



PI Saito S, Tsuzaki Y, Yanagida N;  
 XX WPI, 1997-503046/46.  
 DR N-PSDB; AAT96595.  
 XX  
 PT Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide - for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 PS Disclosure; Page 16-19; 51pp; Japanese.  
 XX  
 CC This sequence represents the chimeric protein 40 K-S which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein gB fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 XX  
 SQ Sequence 456 AA;  
 Query Match 99.7%; Score 2317; DB 18; Length 456;  
 Best Local Similarity 99.6%; Pred. No. 5.2e-143;  
 Matches 454; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MHYFRNCIFELIVILYIGTNSPSTQNTSRVSVQSESEFYLCPPVGSTVRL 60  
 DB 1 MHYFRNCIFELIVILYIGTNSPSTQNTSRVSVQSESEFYLCPPVGSTVRL 60  
 QY 61 EFGCMSITKQDANPNNGQTOLEARMELTDLINAKMTLASIQDYAKIEASLSAYSEAE 120  
 DB 61 EFGCMSITKQDANPNNGQTOLEARMELTDLINAKMTLASIQDYAKIEASLSAYSEAE 120  
 QY 121 TVNNNINATLEOLKMAKTNLESAINQANTDKTFDNEHNPVLEAYKALKTTLEQATNLE 180  
 DB 121 TVNNNINATLEOLKMAKTNLESAINQANTDKTFDNEHNPVLEAYKALKTTLEQATNLE 180  
 QY 181 GLSSTAYNOIRNNLVLDLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240  
 DB 181 GLSSTAYNOIRNNLVLDLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240  
 QY 241 KTNADALNSFTKVIQNNQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTYWN 300  
 DB 241 KTNADALNSFTKVIQNNQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTYWN 300  
 QY 301 GDEPSRIILANTNSITDVSWIYSLAGTNKYOFSFSNYPSTGYLYFPYKLVKADANNV 360  
 DB 301 GDEPSRIILANTNSITDVSWIYSLAGTNKYOFSFSNYPSTGYLYFPYKLVKADANNV 360  
 QY 361 GLOYKLNNGNVOQVEPATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 DB 361 GLOYKLNNGNVOQVEPATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 QY 421 EGNMNVAPMIGNIYLSNENNADKIPGYRRPGFTL 456  
 DB 421 EGNMNVAPMIGNIYLSNENNADKIPGYRRPGFTL 456  
 RESULT 2  
 ID AAM36051 standard; Protein; 1086 AA.  
 XX AAM36051;  
 AC AAM36051;  
 XX  
 DT 15-UTL-1998 (first entry)  
 XX  
 DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.  
 XX  
 KM Chimeric; Marek's disease virus; outer membrane protein; fusion protein;  
 XX antigen; vaccine; poultry.  
 KM  
 OS Chimeric - Marek's disease gammaherpesvirus.  
 OS Chimeric - Mycoplasma gallisepticum.  
 XX

FH Key Location/Qualifiers  
 FT Region 1..672 /note="derived from Marek's disease virus gB protein"  
 FT Region 693..1086 /note="derived from M. gallisepticum antigen"  
 FT Region /note="derived from M. gallisepticum antigen"  
 XX  
 XX MO9736924-A1.  
 XX  
 XX 09-OCT-1997.  
 XX  
 XX 28-MAR-1997; 97WO-JP01084.  
 XX  
 XX 29-MAR-1996; 96JP-0103548.  
 XX  
 XX (JAPG) NIPPON ZEON KK.  
 XX  
 PI Saito S, Tsuzaki Y, Yanagida N;  
 XX  
 XX WPI, 1997-503046/46.  
 DR N-PSDB; AAT96596.  
 XX  
 PT Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide - for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 PS Disclosure; Page 22-30; 51pp; Japanese.  
 XX  
 CC This sequence represents the chimeric protein 40 K-C which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein gB fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 XX  
 SQ Sequence 1086 AA;  
 Query Match 86.2%; Score 2004; DB 18; Length 1086;  
 Best Local Similarity 99.2%; Pred. No. 4.2e-122;  
 Matches 393; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 61 EFGCMSITKQDANPNNGQTOLEARMELTDLINAKMTLASIQDYAKIEASLSAYSEAE 120  
 DB 61 EFGCMSITKQDANPNNGQTOLEARMELTDLINAKMTLASIQDYAKIEASLSAYSEAE 120  
 QY 121 TVNNNINATLEOLKMAKTNLESAINQANTDKTFDNEHNPVLEAYKALKTTLEQATNLE 180  
 DB 121 TVNNNINATLEOLKMAKTNLESAINQANTDKTFDNEHNPVLEAYKALKTTLEQATNLE 180  
 QY 181 GLSSTAYNOIRNNLVLDLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240  
 DB 181 GLSSTAYNOIRNNLVLDLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240  
 QY 241 KTNADALNSFTKVIQNNQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTYWN 300  
 DB 241 KTNADALNSFTKVIQNNQSFVGTFTNANVOPSNYSFVAFSADVTPVNYKYARRTYWN 300  
 QY 301 GDEPSRIILANTNSITDVSWIYSLAGTNKYOFSFSNYPSTGYLYFPYKLVKADANNV 360  
 DB 301 GDEPSRIILANTNSITDVSWIYSLAGTNKYOFSFSNYPSTGYLYFPYKLVKADANNV 360  
 QY 361 GLOYKLNNGNVOQVEPATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 DB 361 GLOYKLNNGNVOQVEPATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTG 420  
 QY 421 EGNMNVAPMIGNIYLSNENNADKIPGYRRPGFTL 456  
 DB 1051 EGNMNVAPMIGNIYLSNENNADKIPGYRRPGFTL 1086  
 RESULT 3  
 ID AAR63230 standard; Protein; 615 AA.  
 XX AAR63230  
 XX

AA63230;  
 25-MAR-2003 (updated)  
 23-JUN-1995 (first entry)  
 Mycoplasma gallisepticum antigen (UM-67).  
 recombinant avipox virus; live vaccine; mycoplasma antigen.  
 Mycoplasma gallisepticum.  
 Key Location/Qualifiers  
 Protein 1..615  
 /note= "Trp residues correspond to TGA codons"  
 MO9423019-A1.  
 13-OCT-1994.  
 31-MAR-1994; 94WO-JP00541.  
 31-MAR-1993; 93JP-0074139.  
 30-SEP-1993; 93JP-0245625.  
 (JAPG) NIPPON ZEON KK.  
 (SHIO) SHIONOGI & CO LTD.  
 Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I,  
 Saeki S, Saitoh, Takahashi K;  
 MPI; 1994-333181/41.  
 N-PSDB; AAQ77857.  
 Recombinant avipox virus combining DNA encoding a polypeptide -  
 exhibiting antigenicity of mycoplasma, useful for the production  
 of a live vaccine  
 Claim 4; Page 87-91; 123pp; Japanese.  
 A restriction fragment of the insert of M.gallisepticum genomic clone  
 pUM-67 containing an open reading frame was sequenced (AAQ77857). The  
 ORF encodes an antigenic polypeptide (AA63230). A recombinant avipox  
 virus comprising the coding sequence can be used as a live vaccine to  
 protect against infection by Mycoplasma gallisepticum.  
 (Updated on 25-MAR-2003 to correct PN field.)  
 Sequence 615 AA;  
 Query Match 82.5%; Score 1917; DB 15; Length 615;  
 Best Local Similarity 98.7%; Pred. No. 9.2e-117;  
 Matches 378; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

64 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 123  
 27 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 86  
 124 NNUNATLEOLKMAKTNLESAINQANTDCTPDNEHPNVKALKTLEORATNLEGLS 183  
 87 NNUNATLEOLKMAKTNLESAINQANTDCTPDNEHPNVKALKTLEORATNLEGLA 146  
 184 STAYNOIRNNLVLDYNNKASSLITKTLDPINGGTLDSNEITVNRNINNTLSTINEOKTN 243  
 147 STAYNOIRNNLVLDYNNKASSLITKTLDPINGGTLDSNEITVNRNINNTLSTINEOKTN 206  
 244 ADALSNSTIKKYIONNEOSFVGTFTNANVOPSNSYFVAFSADVTPEVNYKVARITVWNGDE 303  
 207 ADALSNSTIKKYIONNEOSFVGTFTNANVOPSNSYFVAFSADVTPEVNYKVARITVWNGDE 266  
 304 PSSRLIANTNSTIDVSWIYSLAGTNTKYOFSNSYSPSTGYLYFPYKLVKADANNVGLQ 363  
 267 PSSRLIANTNSTIDVSWIYSLAGTNTKYOFSNSYSPSTGYLYFPYKLVKADANNVGLQ 326  
 364 YKLANNVQOVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGECN 423

|||||  
 327 YKLANNVQOVEFATSTSANNTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGECN 386  
 424 MNKVAPMIGNITYLSSNNENADKI 446  
 387 MNKVAPMIGNITYLSSNNENADKI 409  
 RESULT 4  
 ID AAR63229 standard; Protein; 610 AA.  
 AAR63229;  
 25-MAR-2003 (updated)  
 23-JUN-1995 (first entry)  
 Mycoplasma gallisepticum antigen (UM-66).  
 recombinant avipox virus; live vaccine; mycoplasma antigen.  
 Mycoplasma gallisepticum.  
 Key Location/Qualifiers  
 Protein 1..610  
 /note= "Trp residues correspond to TGA codons"  
 MO9423019-A1.  
 13-OCT-1994.  
 31-MAR-1994; 94WO-JP00541.  
 31-MAR-1993; 93JP-0074139.  
 30-SEP-1993; 93JP-0245625.  
 (JAPG) NIPPON ZEON KK.  
 (SHIO) SHIONOGI & CO LTD.  
 Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I,  
 Saeki S, Saitoh, Takahashi K;  
 MPI; 1994-333181/41.  
 N-PSDB; AAQ77856.  
 Recombinant avipox virus combining DNA encoding a polypeptide -  
 exhibiting antigenicity of mycoplasma, useful for the production  
 of a live vaccine  
 Claim 4; Page 78-81; 123pp; Japanese.  
 A restriction fragment of the insert of M.gallisepticum genomic clone  
 pUM-66 containing an open reading frame was sequenced (AAQ77856). The  
 ORF encodes an antigenic polypeptide (AAR63229). A recombinant avipox  
 virus comprising the coding sequence can be used as a live vaccine to  
 protect against infection by Mycoplasma gallisepticum.  
 (Updated on 25-MAR-2003 to correct PN field.)  
 Sequence 610 AA;  
 Query Match 79.6%; Score 1849; DB 15; Length 610;  
 Best Local Similarity 95.0%; Pred. No. 2.5e-112;  
 Matches 363; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

64 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 123  
 27 CMSITKQDANNNGQTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 86  
 124 NNUNATLEOLKMAKTNLESAINQANTDCTPDNEHPNVKALKTLEORATNLEGLS 183  
 87 NNUNATLEOLKMAKTNLESAINQANTDCTPDNEHPNVKALKTLEORATNLEGLS 146  
 184 STAYNOIRNNLVLDYNNKASSLITKTLDPINGGTLDSNEITVNRNINNTLSTINEOKTN 243

```

Db      147 STAYNOIRNNVLVDLYNKASSLITKTLDPINGGTLIDSLDSEITTTANKKINNTLSTINEQKTN 206
QY      244 ADALNSFTKRVIONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 303
Db      207 ADALNSFTKRVIONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 266
QY      304 PSSRIANTNSITDVSWMISLAGTNTKYQSFNSYGPSTGYLYFPYKLVKADANNVGLQ 363
Db      267 PSSRIANTNSITDVSWMISLAGTNTKYQSFNSYGPSTGYLYFPYKLVKADANNVGLQ 326
QY      364 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGREN 423
Db      327 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGREN 386
QY      424 MNKVAPMIGNIYISSNENNADK 445
Db      387 MNKVAPMIGNIYISSNENNADK 408

```

## RESULT 5

AA04493  
ID AA04493 standard; Protein; 368 AA.

AA04493;

25-MAR-2003 (updated)  
16-JUN-1994 (first entry)

Mycoplasma gallisepticum 40kD antigen.

Vaccine; mycoplasma infection; poultry; fowl.

Mycoplasma gallisepticum.

Key Location/Qualifiers

FT Misc-difference 262 /note="corresponds to NNN codon in AA053419"

FT Misc-difference 283 /note="corresponds to NNN codon in AA053419"

PN WO9324646-A1.

XX 09-DEC-1993.

PF 28-MAY-1993; 93WO-JP00715.

PR 29-MAY-1992; 92JP-0138819.

PA (JAPG) NIPPON ZEON KK.

PA (SHIO) SHIONOGI & CO LTD.

PI Aoyama S, Fujisawa A, Iritani Y, Ohkawa S, Saito S;

DR WPI; 1993-405837/50.

DR N-PSDB; AA053419.

XX Mycoplasma gallisepticum antigen and DNA coding for it - useful

XX for vaccination of fowl against mycoplasma infections

XX Claim 2; Page 23-26; 37pp; Japanese.

XX The sequence coding for the 40kDa antigen was obtained by PCR

XX amplification of M.gallisepticum genomic DNA. The antigen reacts

XX with Mycoplasma-immune or Mycoplasma-infected serum and can be used

XX as a vaccine to protect fowl from M.gallisepticum infection.

XX (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 368 AA;

```

QY      64 CMSITKKDANPNNGOTOLEAARMETDLINAKAMTLASLOPYAKIEASLSAYSFAETVN 123
Db      27 CMSITKKDANPNNGOTOLEAARMETDLINAKAMTLASLOPYAKIEASLSAYSFAETVN 86
QY      124 NNLANTEBOLKMAKTNLESAINQANTDKTTFDNEHPNLVEYKALKTLEORATNLEGLS 183
Db      87 NNLANTEBOLKMAKTNLESAINQANTDKTTFDNEHPNLVEYKALKTLEORATNLEGLS 146
QY      184 STAYNOIRNNVLVDLYNKASSLITKTLDPINGGTLIDSLDSEITTTANKKINNTLSTINEQKTN 243
Db      147 STAYNOIRNNVLVDLYNKASSLITKTLDPINGGTLIDSLDSEITTTANKKINNTLSTINEQKTN 206
QY      244 ADALNSFTKRVIONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 303
Db      207 ADALNSFTKRVIONNEOSFVGTFTNANQPSNYSFVASADYTPVNYKYARRTWNGDE 266
QY      304 PSSRIANTNSITDVSWMISLAGTNTKYQSFNSYGPSTGYLYFPYKLVKADANNVGLQ 363
Db      267 PSSRIANTNSITDVSWMISLAGTNTKYQSFNSYGPSTGYLYFPYKLVKADANNVGLQ 326
QY      364 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 400
Db      327 YKLNNGNVQVEFATSTSANNTTANPTPAVDEIKVAK 363

```

## RESULT 6

AA063227  
ID AA063227 standard; Protein; 368 AA.

AA063227;

25-MAR-2003 (updated)  
23-JUN-1995 (first entry)

Mycoplasma gallisepticum 40kD antigen.

recombinant avipox virus; live vaccine; mycoplasma 40kD antigen;

TM-1.

OS Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Misc-difference 262 /note="corresponds to a NNN codon"

FT Misc-difference 283 /note="corresponds to a NNN codon"

PN WO9423019-A1.

PF 31-MAR-1994; 94WO-JP00541.

PR 31-MAR-1993; 93JP-0074139.

PR 30-SEP-1993; 93JP-0245625.

XX (JAPG) NIPPON ZEON KK.

PA (SHIO) SHIONOGI & CO LTD.

PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;

DR WPI; 1994-333181/41.

DR N-PSDB; AA077854.

XX Recombinant avipox virus combining DNA encoding a polypeptide -

XX exhibiting antigenicity of mycoplasma, useful for the production

XX of a live vaccine

XX Claim 4; Page 71-74; 123pp; Japanese.

XX The plasmid pUTM-1P contains a sequence (the TM-1 gene) coding



CC for the 40kD antigen of Mycoplasma gallisepticum under the control  
 CC of a synthetic promoter. A 1300 bp restriction fragment containing  
 CC the promoter-ORF sequence was excised and was used in the  
 CC construction of plasmid pN27929-R2. This in turn was involved in the  
 CC construction of a recombinant avipox virus vector comprising the  
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from  
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FPV  
 CC sequences. The recombinant avipox virus is useful as a live vaccine  
 CC to protect against infection by Mycoplasma gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 368 AA;

Query Match 69.7%; Score 1619; DB 15; Length 368;

Best Local Similarity 95.5%; Pred. No. 1.2e-97;

Matches 322; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 64 CMSITKKDANPNNGCOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 123  
 DB 27 CMSITKKDANPNNGCOTOLEAARMELTDLINAKARTLASLODYAKIEASLSAYSEAEVTN 86  
 QY 124 NNUNATLEOLKMAKTNLESAINOANTDKTTPDNEHPNLVEAYKAKTTLLEORATNLEGLS 183  
 DB 87 NNUNATLEOLKMAKTNLESAINOANTDKTTPDNEHPNLVEAYKAKTTLLEORATNLEGLA 146  
 QY 184 STAYNOIRNNLVLDLYNKASLITKTLDPLNGCTLLDSNEITTVARNINNTLSTINEOKTN 243  
 DB 147 STAYNOIRNNLVLDLYNKASLITKTLDPLNGCTLLDSNEITTVARNINNTLSTINEOKTN 206  
 QY 244 ADALSNFIRKVIQNNQOSFVGFTFNANVOPSNSYFVAFSADVTVPVVKYARRTVMNGDE 303  
 DB 207 ADALSNFIRKVIQNNQOSFVGFTFNANVOPSNSYFVAFSADVTVPVVKYARRTVMNGDE 266  
 QY 304 PSSRIILANTNSITDVSMTYISLAGTNTKYOFSNSYGSPTGYLYPYKLVKADANNVGLQ 363  
 DB 267 PSSRIILANTNSITDVSMTYISLAGTNTKYOFSNSYGSPTGYLYPYKLVKADANNVGLQ 326  
 QY 364 YKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKYAK 400  
 DB 327 YKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKYAK 363

RESULT 7

AAR76955 standard; Protein; 368 AA.

XX AAR76955;

XX 25-MAR-2003 (updated)

DT 12-MAR-1996 (first entry)

XX Mycoplasma gallisepticum antigenic protein TTM-1.

XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1.

XX Mycoplasma gallisepticum.

XX Key Location/Qualifiers

FT Misc-difference 262 /note= "any amino acid"

FT Misc-difference 283 /note= "any amino acid"

XX JF07133295-A.

XX 23-MAY-1995.

XX 27-AUG-1993; 93UP-0213102.

XX 27-AUG-1993; 93UP-0213102.

XX (JAPG ) NIPPON ZEON KK.

PA (SHIO ) SHIONOGI & CO LTD.

XX WPI; 1995-220782/29.

DR N-PSDB; AAQ94711.

XX A new antigenic protein which reacts with Mycoplasma gallisepticum -  
 PT is useful in a component vaccine for use against poultry infected  
 XX with M. gallisepticum.

XX Claim 6; Figs 5-6; 33pp; Japanese.

CC AAQ94711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein

CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum

CC infectious diseases in poultry, and as a diagnostic agent for

CC M. gallisepticum.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 368 AA;

Query Match 69.3%; Score 1611; DB 16; Length 368;

Best Local Similarity 95.0%; Pred. No. 4.1e-97;

Matches 320; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 64 CMSITKKDANPNNGCOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVTN 123  
 DB 27 CMSITKKDANPNNGCOTOLEAARMELTDLINAKARTLASLODYAKIEASLSAYSEAEVTN 86  
 QY 124 NNUNATLEOLKMAKTNLESAINOANTDKTTPDNEHPNLVEAYKAKTTLLEORATNLEGLS 183  
 DB 87 NNUNATLEOLKMAKTNLESAINOANTDKTTPDNEHPNLVEAYKAKTTLLEORATNLEGLA 146  
 QY 184 STAYNOIRNNLVLDLYNKASLITKTLDPLNGCTLLDSNEITTVARNINNTLSTINEOKTN 243  
 DB 147 STAYNOIRNNLVLDLYNKASLITKTLDPLNGCTLLDSNEITTVARNINNTLSTINEOKTN 206  
 QY 244 ADALSNFIRKVIQNNQOSFVGFTFNANVOPSNSYFVAFSADVTVPVVKYARRTVMNGDE 303  
 DB 207 ADALSNFIRKVIQNNQOSFVGFTFNANVOPSNSYFVAFSADVTVPVVKYARRTVMNGDE 266  
 QY 304 PSSRIILANTNSITDVSMTYISLAGTNTKYOFSNSYGSPTGYLYPYKLVKADANNVGLQ 363  
 DB 267 PSSRIILANTNSITDVSMTYISLAGTNTKYOFSNSYGSPTGYLYPYKLVKADANNVGLQ 326  
 QY 364 YKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKYAK 400  
 DB 327 YKLNGNVQOQVEFATSTSANNTTANPTPAVDEIKYAK 363

RESULT 8

AAP93646 standard; Protein; 235 AA.

XX AAP93646;

XX 25-MAR-2003 (updated)

DT 11-MAY-1990 (first entry)

XX Amino acid sequence of Mycoplasma gallisepticum (MG1) polypeptide.

XX Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0305441.

XX 31-MAY-1989; 89EP-0305441.

XX 02-JUN-1988; 88UP-0136343.

XX (JAPG ) NIPPON ZEON KK.

PA (SHIO ) SHIONOGI SEIYAKU KK.

PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI SEIYAKU KK.  
 XX  
 PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;  
 DR WPI, 1989-358393/49.  
 DR N-PSDB; AAN92568.  
 XX  
 PT New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
 PT vaccines.  
 PS  
 XX Disclosure; Fig.1a; 31pp; English.  
 CC This amino acid sequence of MG1 is encoded by M1 DNA and elicits an  
 CC antigen-antibody reaction with anti-MG poultry sera. It can be used as a  
 CC vaccine to prevent and diagnose MG infection. Doseage is at least 1 micro  
 CC gram vaccine / kg. No acute toxicity was noted with a dose of 5 mg / kg.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 235 AA;  
 Query Match 49.4%; Score 1149; DB 10; Length 235;  
 Best Local Similarity 97.9%; Pred. No. 2.9e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 64 CMSITKKDANPNNGOTOLEAARMELTDLINAKMTLASIQDYAKIEASISAYSEAEVTN 123  
 DB 1 CMSITKKDANPNNGOTOLEAARMELTDLINAKMTLASIQDYAKIEASISAYSEAEVTN 60  
 QY 124 NNLAATLEQLKMAKTNLESAINQANTDKTFPDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
 DB 61 NNLAATLEQLKMAKTNLESAINQANTDKTFPDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 QY 184 STAVNQIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 121 STAVNQIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 180  
 QY 244 ADALSNSTFKKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTV 298  
 DB 181 ADALSNSTFKKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTV 235  
 RESULT 9  
 AAR05081  
 ID AAR05081 standard; protein; 235 AA.  
 XX  
 AC AAR05081;  
 XX  
 DT 10-MAR-2003 (updated)  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE MG-1 antigen.  
 XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.  
 OS Mycoplasma gallisepticum.  
 XX  
 PN JP02111795-A.  
 XX  
 PD 24-APR-1990.  
 XX  
 PF 02-JUN-1989; 89JP-0136343.  
 XX  
 PR 02-JUN-1989; 89JP-0136343.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI KK.  
 XX  
 DR WPI, 1990-169109/22.  
 DR N-PSDB; AAQ04686.

XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilizes  
 PT antigen protein of the disease and recombinant vector  
 PT incorporated with its coding gene.  
 PS  
 XX Claim 2; Fig 1a; 20pp; Japanese.  
 CC DNA encoding the protein can be inserted into an expression vector  
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 CC be ligated to other DNA to produce fusion proteins with an N-terminal  
 CC bacterial enzyme sequence.  
 CC See also AAR05081-2 and AAR06437-41.  
 CC (Updated on 10-MAR-2003 to add missing OS field.)  
 XX  
 SQ Sequence 235 AA;  
 Query Match 49.4%; Score 1149; DB 11; Length 235;  
 Best Local Similarity 97.9%; Pred. No. 2.9e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 64 CMSITKKDANPNNGOTOLEAARMELTDLINAKMTLASIQDYAKIEASISAYSEAEVTN 123  
 DB 1 CMSITKKDANPNNGOTOLEAARMELTDLINAKMTLASIQDYAKIEASISAYSEAEVTN 60  
 QY 124 NNLAATLEQLKMAKTNLESAINQANTDKTFPDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
 DB 61 NNLAATLEQLKMAKTNLESAINQANTDKTFPDNEHPNLVEAYKALKTTLEQRATNLEGLS 120  
 QY 184 STAVNQIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 121 STAVNQIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 180  
 QY 244 ADALSNSTFKKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTV 298  
 DB 181 ADALSNSTFKKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTV 235  
 RESULT 10  
 AAR05082  
 ID AAR05082 standard; protein; 261 AA.  
 XX  
 AC AAR05082;  
 XX  
 DT 10-MAR-2003 (updated)  
 DT 08-OCT-1990 (first entry)  
 XX  
 DE TMG-1 antigen.  
 XX  
 KW Mycoplasma gallisepticum; poultry; vaccine.  
 OS Mycoplasma gallisepticum.  
 XX  
 PN JP02111795-A.  
 XX  
 PD 24-APR-1990.  
 XX  
 PF 02-JUN-1989; 89JP-0136343.  
 XX  
 PR 02-JUN-1989; 89JP-0136343.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI KK.  
 XX  
 DR WPI, 1990-169109/22.  
 DR N-PSDB; AAQ04687.  
 XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilizes  
 PT antigen protein of the disease and recombinant vector  
 PT incorporated with its coding gene.  
 PS  
 XX Claim 2; Fig 2; 20pp; Japanese.

CC DNA encoding the protein can be inserted into an expression vector  
 CC for the prodn. of MG-1 polypeptide which elicits an antigen-antibody  
 CC reaction with anti-mycoplasma gallisepticum poultry sera. It may also  
 CC be ligated to other DNA to produce fusion proteins with an N-terminal  
 CC bacterial enzyme sequence.  
 CC See also AAR05081 and AAR06437-41.  
 CC (Updated on 10-MAR-2003 to add missing OS field.)

XX Sequence 261 AA;

Query Match 49.4%; Score 1149; DB 11; Length 261;  
 Best Local Similarity 97.9%; Pred. No. 3,4e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 64 CMSGTKDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVYN 123  
 DB 27 CMSGITKQDANPNNGOTOLEAARMELTDLINAKARTLASLODYAKIEASLSAYSEAEVYN 86  
 QY 124 NNLAATLEQLKMAKTNIESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
 DB 87 NNLAATLEQLKMAKTNIESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
 QY 184 STANQIRNNLVLDYNNASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 147 STANQIRNNLVLDYNNASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 206  
 QY 244 ADALSNFIRKVIQNNESFVGFTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 298  
 DB 207 ADALSNFIRKVIQNNESFVGFTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 261

RESULT 11  
 AAR79911

ID AAR79911 standard; Protein; 261 AA.

AC AAR79911;

DT 19-JUL-1996 (first entry)

DE M.gallisepticum 261 amino acid protein.

KW Detection; probe; primer; PCR; amplification; secretion; lung;  
 KM avian chronic respiratory disease; respiratory tract; nasal cavity.

XX Mycoplasma gallisepticum.

OS JF07236498-A.

PN 12-SEP-1995.

PD 25-FEB-1994; 94JP-0052764.

PF 25-FEB-1994; 94JP-0052764.

PR (JAPG ) NIPPON ZEON KK.

PA (SHIO ) SHIONOGI & CO LTD.

XX WPI; 1995-347462/45.

DR N-PSDB; AAT04076.

XX Detection of Mycoplasma gallisepticum - for the quick detection,  
 PT i.e. within one day, of avian chronic respiratory diseases

XX Claim 3; Page 10-11; 11pp; Japanese.

XX This is the amino acid sequence of a 261 amino acid protein encoded  
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding  
 CC sequence and the sequence of AAT04075 (encoding a 661 amino acid  
 CC protein) can be used to detect M.gallisepticum using probes based on  
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the  
 CC complement of bases 893-919, 1908-1934 and the complement of bases  
 CC 214-2210 of AAT04075 and a probe based on nucleotides 718-41 of the DNA  
 CC encoding this protein. The method using these sequences is faster i.e. is

CC able to detect M.gallisepticum, which causes avian chronic respiratory  
 CC diseases, within one day, from avian secretions, washings from the lung,  
 CC respiratory tract, nasal cavity, etc.

XX Sequence 261 AA;

Query Match 49.4%; Score 1149; DB 16; Length 261;  
 Best Local Similarity 97.9%; Pred. No. 3,4e-67;  
 Matches 230; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 64 CMSGTKDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEVYN 123  
 DB 27 CMSGITKQDANPNNGOTOLEAARMELTDLINAKARTLASLODYAKIEASLSAYSEAEVYN 86  
 QY 124 NNLAATLEQLKMAKTNIESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
 DB 87 NNLAATLEQLKMAKTNIESAINOANTDKTTFDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
 QY 184 STANQIRNNLVLDYNNASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 243  
 DB 147 STANQIRNNLVLDYNNASSLITKTLDPNGTLLDSNEITTVNRNINNTLSTINEOKTN 206  
 QY 244 ADALSNFIRKVIQNNESFVGFTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 298  
 DB 207 ADALSNFIRKVIQNNESFVGFTFTNANVOPSNYSFVAFSADVTPVNYKYARRTV 261

RESULT 12  
 AAP93959

ID AAP93959 standard; Protein; 261 AA.

AC AAP93959;

DT 25-MAR-2003 (updated)

DT 11-MAY-1990 (first entry)

DE Amino acid (AA) sequence of TWG-1 polypeptide.

KW Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

OS EP345021-A.

PN 06-DEC-1989.

PD 31-MAY-1989; 89EP-0305441.

PF 31-MAY-1989; 89EP-0305441.

PR 02-JUN-1988; 88JP-0136343.

PA (JAPG ) NIPPON ZEON KK.

PA (SHIO ) SHIONOGI SEIYAKU KK.

XX (JAPG ) NIPPON ZEON KK.

XX (SHIO ) SHIONOGI SEIYAKU KK.

PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

DR WPI; 1989-358393/49.

DR N-PSDB; AAN92574.

XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry  
 PT vaccines.

XX Disclousure; Fig.2; 31pp; English.

XX This AA sequence of TWG-1 is encoded by TW-1 base sequence. It has the  
 CC same sequence as that of a polypeptide expressed in Mycoplasma  
 CC gallisepticum in nature. When the corresponding DNA sequence is inserted  
 CC into a recombinant vector used to transform a host the antigen protein  
 CC produced can be used as a vaccine to prevent and diagnose MG infection.  
 CC (Updated on 25-MAR-2003 to correct PF field.)

CC (Updated on 25-MAR-2003 to correct PA field.)  
CC (Updated on 25-MAR-2003 to correct PI field.)  
XX  
SQ Sequence 261 AA;  
Query Match 48.4%; Score 1124; DB 10; Length 261;  
Best Local Similarity 97.0%; Pred. No. 1.4e-65;  
Matches 229; Conservative 2; Mismatches 3; Indels 2; Gaps 2;  
QY 64 CMSTIRKQANPNNGQTOLEAAMELTDLINAKAMTLASIQDYAKIEASISASVSAETYN 123  
DB 27 CMSTIRKQANPNNGQTOLEAAMELTDLINAKARLASIQDYAKIEASISASVSAETYN 86  
QY 124 NNUNATLEOLKMAKTNLESAINQANTDKTTPNEHPNLVEAYKALKTTLEQATNLEGS 183  
DB 87 NNUNATLEOLKMAKTNLESAINQANTDKTTPNEHPNLVEAYKALKTTLEQATNLEGS 146  
QY 184 STAYNOIRNNLVLDYNNKASLITTKTLDPNGSTLLDSNEITTVNRRNINTSTINE-OXT 242  
DB 147 STAYNOIRNNLVLDYNNKASLITTKTLDPNGSTLLDSNEITTVNRRNINTSTINE-OXT 205  
QY 243 NADALSNSEFIKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTPVNYKARRTV 298  
DB 206 NADALSNSEFIKVIQNNESFVGTFTNANVOPSNYSFVAFSADVTPVNYKARRTV 261  
RESULT 13  
AAR63226  
ID AAR63226 standard; Protein; 661 AA.  
XX  
AC AAR63226;  
XX  
DT 25-MAR-2003 (updated)  
DT 23-JUN-1995 (first entry)  
XX  
DE Mycoplasma gallisepticum antigen (UM-81).  
XX  
KM recombinant avipox virus; live vaccine; mycoplasma antigen.  
XX  
OS Mycoplasma gallisepticum.  
XX  
FH Key Location/Qualifiers  
FT Protein 1..661  
FT /note= "Tyrp residues correspond to TGA codons"  
XX  
PN W09423019-A1.  
XX  
PD 13-OCT-1994.  
XX  
PF 31-MAR-1994; 94MO-JP00541.  
XX  
PR 31-MAR-1993; 93JP-0074139.  
PR 30-SEP-1993; 93JP-0245625.  
XX  
PA (JAPG ) NIPPON ZEON KK.  
PA (SHIO ) SHIONOGI & CO LTD.  
XX  
PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohnawa T;  
PI Saeki S, Saitos, Takahashi K;  
XX  
DR WPI; 1994-333181/41.  
DR N-PSDB; AAQ77853.  
XX  
PT Recombinant avipox virus combining DNA encoding a polypeptide -  
PT exhibiting antigenicity of mycoplasma, useful for the production  
PT of a live vaccine  
XX  
PS Claim 4; Page 61-65; 123pp; Japanese.  
XX  
CC A restriction fragment of the insert of M.gallisepticum genomic clone  
CC PUM-81 containing an open reading frame was sequenced (AAQ77853). The  
CC ORF encodes an antigenic polypeptide (AAR63226). A recombinant avipox  
CC virus comprising the coding sequence can be used as a live vaccine to

CC Protect against infection by Mycoplasma gallisepticum.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 661 AA;  
Query Match 34.6%; Score 803; DB 15; Length 661;  
Best Local Similarity 41.4%; Pred. No. 4.2e-44;  
Matches 187; Conservative 64; Mismatches 129; Indels 72; Gaps 13;  
QY 51 PPVGSIVRLRFGCMSTIRKQANPNNGQTOLEAAMELTDLINAKAMTLASIQ 103  
DB 35 PPVGSIVRLRFGCMSTIRKQANPNNGQTOLEAAMELTDLINAKAMTLASIQ 84  
QY 104 DYAKIEASISASVSAETYNUNNATLEOLKMAKTNLESAINQANTDKTTPNEHPNLVE 163  
DB 85 DYAKIEASISASVSAETYNUNNATLEOLKMAKTNLESAINQANTDKTTPNEHPNLVE 144  
QY 164 AYKALKTTLEQATNLEGSSTAYNOIRNNLVLDYNNKASLITTKTLDPNGSTLLDSNEI 223  
DB 145 AYKALKTTLEQATNLEGSSTAYNOIRNNLVLDYNNKASLITTKTLDPNGSTLLDSNEI 203  
QY 224 TTVNRRNINTSTINEOQTNADALSNSEFIKVIQNNESFVGTFTNANVOPSNYSFVAFS 283  
DB 204 TTVNRRNINTSTINEOQTNADALSNSEFIKVIQNNESFVGTFTNANVOPSNYSFVAFS 262  
QY 284 ADVTPVNYKARRTVYNNNGDE-----PESRIIANTNSITDVSWIYSLAGTNTKQFSPSN 337  
DB 263 VNVDTPNMNFQKRYKVASSENTPLATTTPAEDTQQAASLITDVSWIYSLAGTNTKQFSPSN 322  
QY 338 YG-PSTGYLPPPYKLVKADANNVGLQYKLNNGVQVEF-----ATSTSAN--NTTA 387  
DB 323 FGAERTALVLPYKLVKXSD--NVGLQYKLNNGVQVEF-----ATSTSAN--NTTA 380  
QY 388 NP-----TPAVDEIKVAKIVISGRFGQNTIEISVPTGEGMNNKVAIPGIGIYSS---- 438  
DB 381 SPAEQSAPVTDIDIKIAKVALNALNKFNSNTIEFSVPTG-----KAAPIGMNMYLTSNSE 435  
QY 439 -----NENN-----ADKIPGV 449  
DB 436 VNRKKIYDLDLFGNSFNENNPTAVTVLDLKKY 467  
RESULT 14  
AAR79910  
ID AAR79910 standard; Protein; 661 AA.  
XX  
AC AAR79910;  
XX  
DT 19-JUL-1996 (first entry)  
XX  
DE M.gallisepticum 661 amino acid protein.  
XX  
KM Detection; probe; primer; PCR; amplification; secretion; lung;  
KM avian chronic respiratory disease; respiratory tract; nasal cavity.  
XX  
OS Mycoplasma gallisepticum.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 220  
FT /note= "encoded by TGA"  
FT Misc-difference 270  
FT /note= "encoded by TGA"  
FT Misc-difference 305  
FT /note= "encoded by TGA"  
FT Misc-difference 581  
FT /note= "encoded by TGA"  
XX  
PN JP07236498-A.  
XX  
PD 12-SEP-1995.  
XX  
PR 25-FEB-1994; 94JP-0052764.

PR 25-FEB-1994; 94JP-0052764.  
 XX (JAP) NIPPON ZEON KK.  
 PA (SHIO) SHIONOGI & CO LTD.  
 XX  
 DR WPI; 1995-347462/45.  
 DR N-PSDB; AAT04075.  
 XX  
 PT Detection of Mycoplasma gallisepticum - for the quick detection,  
 XX i.e. within one day, of avian chronic respiratory diseases  
 PS Claim 2; Page 7-10; 11pp; Japanese.  
 XX  
 CC This is the amino acid sequence of a 661 amino acid protein encoded  
 CC by a fragment of the Mycoplasma gallisepticum genome. The encoding  
 CC sequence and the sequence of AAT04076 (encoding a 261 amino acid  
 CC protein) can be used to detect M.gallisepticum using probes based on  
 CC nucleotides 1125-1648 and primers based on nucleotides 449-466, the  
 CC complement of bases 893-919, 1908-1934 and the complement of bases  
 CC 2184-2210 of the sequence, and a probe based on nucleotides 718-41 of  
 CC AAT04076. The method using these sequences is faster i.e. is able to  
 CC detect M.gallisepticum, which causes avian chronic respiratory diseases,  
 CC within one day, from avian secretions, washings from the lung,  
 CC respiratory tract, nasal cavity, etc.  
 CC  
 XX Sequence 661 AA;  
 SQ  
 Query Match 34.6%; Score 803; DB 16; Length 661;  
 Best Local Similarity 41.4%; Fred. No. 4.2e-44;  
 Matches 187; Conservative 64; Mismatches 129; Indels 72; Gaps 13;  
 QY 51 PVGVSTVIRLEFGCGMSITKKDANPNNGQTOLEAARMEITD-----LINAKAMTLASIQ 103  
 DB 35 PPSGG-----NNNGCNTNPBGCGMMNAAKELADAKAALTITLINGETANLASYE 84  
 QY 104 DYAKIEASLSAYSEAEVTNNNNLNTLEQLMAKTNLESAINQANTDKTTFDNEHNLVE 163  
 DB 85 DYAKIKSELTSAVETAKAVSAKTGATITNEVNEAKTTLDAIKKASANDPDAQHGLIVE 144  
 QY 164 AKYKAKTLLEQRATNLEGLSTAYNQIRNNIVDLYNKASSLITKTLIDLNGTILDSBEI 223  
 DB 145 AYNNIKETLKEKENTLDSLANNENVAIITNINSLYEKANTIVTATLDEAT-GNIPVMSV 203  
 QY 224 TVVRNINNTLSTINEOKTNADALNSPFKVIQNNESPFGTFTNANVQPSNYSFVAFS 283  
 DB 204 TOANODITNATSLRLAKQNDNLANSITKQSLVKNMTRDVANNQ-QPANISFVGF 262  
 QY 284 ADVTPVNYKVARRTVWNGD-----PSSRLANTNSITDVSWIYSLAGNTKYQSPFN 337  
 DB 263 VNVDPNPNFPAQRKVMASENTPLATTPAEDATQQAASLTDVSWIYSLNGAEKATLSFRY 322  
 QY 338 YG-BSYGLVFFPKYKAAADANNVGLYKLNANVQVEF-----ATSTSAN-NTTA 387  
 DB 323 FGAERTALVFFPKYKVTSD--NVGLQYKLNAGDTRQINFVQTPASGSSDVAANEERTMA 380  
 QY 388 NP-----TPADEVIVAKIVISGLRFGQNTIELSVPTGEGMKNKAPMIGNITVLS 438  
 DB 381 SPAEQSAPTYDDIKIAKVALSNLKFNSNTIEFSVPTG-----KAPMIGMNTLSSNSE 435  
 QY 439 -----NENN-----ADKIPGY 449  
 DB 436 VNKKKIYDPLFGNSFNENNPPAVLVLDLKG 467  
 RESULT 15  
 AAM11978  
 ID AAM11978 standard; Protein; 647 AA.  
 XX  
 AC AAM11978;  
 XX  
 DT 23-APR-1997 (first entry)  
 XX  
 DE Mycobacterium gallisepticum pmGAL.2 adhesin.

XX  
 KW Adhesin; pmGAL; mycoplasma; diagnosis; vaccine; vector;  
 KM respiratory disease; poultry; haemagglutinin.  
 XX  
 OS Mycoplasma gallisepticum strain S6.  
 XX  
 FH Key  
 FT Peptide  
 FT  
 FT Location/Qualifiers  
 FT 1..25  
 FT /label= sig peptide  
 FT /note= "the signal peptide shows homology to  
 FT the pmGAL.3 signal peptide"  
 XX  
 XX CA2135330-A.  
 XX  
 XX 11-MAY-1995.  
 XX  
 XX 08-NOV-1994; 94CA-2135330.  
 XX  
 XX 20-APR-1994; 94US-0230312.  
 XX 10-NOV-1993; 93AU-0050593.  
 XX  
 PA (BROW/) BROWNING G F.  
 XX  
 XX Browning GF, Glew MD, Markham PF, Walker ID, Whitehear KG;  
 XX  
 XX WPI; 1995-241027/32.  
 XX  
 DR N-PSDB; AAT51531.  
 XX  
 XX New promoter region from a Mycoplasma gallisepticum adhesin gene -  
 XX useful when coupled to foreign antigen gene, for prodn. of  
 XX and manipulating live vaccines, also new probes for detecting Mycoplasma  
 XX  
 XX Disclosure; Fig 3; 81pp; English.  
 XX  
 CC Adhesin pmGAL.2 (AAM11978) and adhesin pmGAL.3 fragment (AAM11979)  
 CC are products of gene sequences (see also AAT51531) isolated from  
 CC Mycoplasma gallisepticum. DNA constructs incorporating the  
 CC promoter and/or signal sequences of the pmGA gene can be used  
 CC in the prodn. of multivalent live vaccines. The signal peptide  
 CC sequence is utilised where attachment of an exogenous antigen  
 CC gene to the mycoplasma cell membrane is required.  
 CC  
 XX  
 SQ Sequence 647 AA;  
 Query Match 30.6%; Score 711.5; DB 16; Length 647;  
 Best Local Similarity 39.0%; Fred. No. 3.8e-38;  
 Matches 173; Conservative 63; Mismatches 135; Indels 73; Gaps 13;  
 QY 53 VGVSTVIRLEFGCGMSITKK-----DANPNNGQ-----TOLEAARMEITD 90  
 DB 15 IGSFVMLAASCTTPTNPPTPNPPPSGCGMNGDTNPGDGGMMNAAQELAAARMLTT 74  
 QY 91 LINAKAMTLASLDYAKIEASLSAYSEAEVTNNNNLNTLEQLMAKTNLESAINQANTD 150  
 DB 75 IFDSKANLGLVYDVKKTQNTLTKAYDAKTVLNNSSSTQKLNKAEKTRLETARTATS 134  
 QY 151 KTTFDNEHNPUNVEYKAKLTKTLLEQRATNLEGLSTAYNQIRNNIVDLYNKASSLITKTL 210  
 DB 135 KQTFDEQHAELVYKVKELKTTLSNETATLAPYADAQYAGIMHLSGLYDACKAATTKYLE 194  
 QY 211 PLNGTILDSNEITVRNINNTL--STINEOKTN-----ADALNSPFKVIQNNESFV 264  
 DB 195 PVEGDP-LTSAVMMANTKIYEAIKDEVLDNOKENAKTKLADSLSSIVKKTGYBE----- 249  
 QY 265 GTFTNANVQPSNYSFVAFSADVTPV-----NYKYARRTW-NGDEPSSRLIANT--- 312  
 DB 250 ---AHNKAQPANYSFVGQKRWYTELDDKQVFPWMDYABERTIFNPSDEP-RSISNTPAD 304  
 QY 313 -----NSITDVSWIYSLAGNTKYQSPFNYSYGLVFFPKYKAAADANNVGLYKLN 367  
 DB 305 GQTMAPLSNYSWISYSLAGTAKYTLLEFTYGGPSYGLVFFPKYKLVNTSDQVGLYKLN 364

Qy	368	NGWQOEFATSTSA-----NNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEG	422
Db	365	D-----ATKPSAITTFGSDQTMNGKTFPTVNDINVAKVITLANLFGSNKIEFSVPA---	413
Qy	423	NNMKVAPMIGNIYLSSENENNADKI	446
		:     :    :	
Db	414	--EKVSPMIGNMYLSSSPNNMNKI	435

Search completed: August 14, 2003, 10:21:07  
Job time : 87 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 10:19:47 ; Search time 29 Seconds

(without alignments)  
665,302 Million cell updates/sec

Title: US-09-901-572A-3

Perfect score: 2324  
Sequence: 1 MHYERRNCIFFLVILYGTN.....SSNNENADKIPGRRPPTFL 456

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfltest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1917	82.5	615	2	US-08-525-742-10 Sequence 10, Appli
2	1849	79.6	610	2	US-08-525-742-8 Sequence 8, Appli
3	1645	70.8	368	2	US-08-525-742-4 Sequence 4, Appli
4	1619	69.7	368	1	US-08-185-851A-4 Sequence 4, Appli
5	803	34.6	661	2	US-08-525-742-2 Sequence 2, Appli
6	314	13.5	865	1	US-07-803-633A-13 Sequence 13, Appli
7	181.5	7.8	10182	4	US-09-134-001C-3159 Sequence 3159, Ap
8	176.5	7.6	933	3	US-08-293-728-2 Sequence 2, Appli
9	176.5	7.6	933	3	US-09-421-868-2 Sequence 4, Appli
10	174.5	7.5	1041	1	US-08-220-151-4 Sequence 4, Appli
11	174.5	7.5	1041	1	US-08-413-118-4 Sequence 4, Appli
12	174.5	7.5	1041	1	US-08-473-446-4 Sequence 4, Appli
13	150.5	6.5	682	4	US-08-836-687B-37 Sequence 37, Appli
14	147.5	6.3	1073	3	US-09-541-782-6 Sequence 6, Appli
15	147.5	6.3	1073	3	US-09-723-820-6 Sequence 6, Appli
16	139.5	6.0	930	4	US-09-134-001C-5314 Sequence 5314, Ap
17	139	6.0	1004	4	US-09-268-347-30 Sequence 30, Appli
18	137.5	5.9	1002	4	US-09-268-347-24 Sequence 24, Appli
19	137.5	5.9	2137	4	US-09-134-001C-4463 Sequence 4463, Ap
20	137	5.9	1098	3	US-08-923-992A-8 Sequence 8, Appli
21	135	5.8	943	3	US-08-911-321-4 Sequence 4, Appli
22	134.5	5.8	1073	3	US-09-206-942-49 Sequence 49, Appli
23	134.5	5.8	1079	4	US-09-206-942-47 Sequence 47, Appli
24	133.5	5.7	1164	3	US-08-923-992A-2 Sequence 2, Appli
25	132	5.7	1104	3	US-08-923-992A-4 Sequence 22, Appli
26	129.5	5.6	903	3	US-08-804-439A-22 Sequence 22, Appli
27	129.5	5.6	903	3	US-08-720-229-22 Sequence 22, Appli

28	129.5	5.6	904	3	US-08-632-537-1 Sequence 1, Appli
29	129.5	5.6	904	5	PCT-US96-05316-1 Sequence 1, Appli
30	129.5	5.6	904	6	5244792-4 Sequence 4, Appli
31	129	5.6	1833	4	US-08-621-944A-4 Sequence 4, Appli
32	129	5.6	1833	4	US-08-945-567D-4 Sequence 4, Appli
33	129	5.6	1992	4	US-08-621-944A-3 Sequence 3, Appli
34	129	5.6	1992	4	US-08-945-567D-3 Sequence 3, Appli
35	129	5.6	2048	4	US-09-268-347-48 Sequence 48, Appli
36	128.5	5.5	903	1	US-08-220-151-8 Sequence 8, Appli
37	128.5	5.5	903	1	US-08-413-118-8 Sequence 8, Appli
38	128.5	5.5	903	3	US-08-473-446-8 Sequence 8, Appli
39	128.5	5.5	1095	4	US-09-206-942-45 Sequence 45, Appli
40	128.5	5.5	1101	4	US-09-206-942-43 Sequence 43, Appli
41	128	5.5	682	3	US-08-481-435-6 Sequence 6, Appli
42	127.5	5.5	1128	3	US-08-923-992A-6 Sequence 6, Appli
43	127.5	5.5	1164	3	US-08-923-992A-10 Sequence 10, Appli
44	127	5.5	891	1	US-08-042-747A-6 Sequence 6, Appli
45	126	5.4	666	3	US-08-961-083-2 Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-525-742-10  
Sequence 10, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Sasaki, Sakiko  
APPLICANT: Ohkawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
APPLICANT: Takahashi, Kiyoohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEROFO  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESSER: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525, 742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930

TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 615 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-10

Query Match 82.5%; Score 1917; DB 2; Length 615;  
Best Local Similarity 98.7%; Pred. No. 1,5e-126;  
Matches 378; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 64 CMSITKQDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEYVN 123  
DB 27 CMSITKQDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEYVN 86  
QY 124 NNLANATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
DB 87 NNLANATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEQRATNLEGLA 146  
QY 184 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 206  
QY 244 ADALNSFTKVIQNNNEOSFVGFTPTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 303  
DB 207 ADALNSFTKVIQNNNEOSFVGFTPTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 266  
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYPPYKLVKADANNVGLQ 363  
DB 267 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYPPYKLVKADANNVGLQ 326  
QY 364 YKLNNGNVQVEFAFSTGSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 423  
DB 327 YKLNNGNVQVEFAFSTGSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386  
QY 424 MNKVAPMIGNIYLSNENNAADKI 446  
DB 387 MNKVAPMIGNIYLSNENNAADKI 409

## RESULT 2

US-08-525-742-8  
Sequence 8, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Onosawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-8

Query Match 79.6%; Score 1849; DB 2; Length 610;  
Best Local Similarity 95.0%; Pred. No. 8.7e-122;  
Matches 363; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 64 CMSITKQDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEYVN 123  
DB 27 CMSITKQDANPNNGOTOLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSEAEYVN 86  
QY 124 NNLANATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEQRATNLEGLS 183  
DB 87 NNLANATLEQLKMAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTTLEQRATNLEGLS 146  
QY 184 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKASSLITKTLDPINGGTLDSNEITTVNRNINNTLSTINEOKTN 206  
QY 244 ADALNSFTKVIQNNNEOSFVGFTPTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 303  
DB 207 ADALNSFTKVIQNNNEOSFVGFTPTNANVQPSNYSFVAFSADVTVPVNYKVARRTVMNGDE 266  
QY 304 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYPPYKLVKADANNVGLQ 363  
DB 267 PSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYPPYKLVKADANNVGLQ 326  
QY 364 YKLNNGNVQVEFAFSTGSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 423  
DB 327 YKLNNGNVQVEFAFSTGSANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGN 386  
QY 424 MNKVAPMIGNIYLSNENNAADKI 445  
DB 387 MNKVAPMIGNIYLSNENNAADKI 408

## RESULT 3

US-08-525-742-4  
Sequence 4, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Onosawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iritani, Yoshikazu  
APPLICANT: Aoyama, Shigemi



APPLICANT: Takahashi, Kiyochito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
TITLE OF INVENTION: AS USE THEREOF  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESS: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESS: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-4

Query Match 70.8%; Score 1645; DB 2; Length 368;  
Best Local Similarity 96.1%; Pred. No. 8.7e-108;  
Matches 324; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 64 CMSITKQDANPNNGQTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSAEATVN 123  
DB 27 CMSITKQDANPNNGQTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSAEATVN 86  
QY 124 NNLAATLEQLMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLA 183  
DB 87 NNLAATLEQLMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLA 146  
QY 184 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEOKTN 206  
QY 244 ADALNSFIKKVYQNNQSFVGTFTNANVQPSNYSFVAFSADVTVPVYKVARRTVWNGDE 303  
DB 207 ADALNSFIKKVYQNNQSFVGTFTNANVQPSNYSFVAFSADVTVPVYKVARRTVWNGDE 266  
QY 304 PSSRLIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKADANNVGLQ 363  
DB 267 PSSRLIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKADANNVGLQ 326  
QY 364 YKLNGNVQVEFATSTSNANTTANPTPAVDEIKVAK 400

DB 327 YKLNGNVQVEFATSTSNANTTANPTQQLMRKLK 363

RESULT 4  
US-08-185-851A-4  
Sequence 4, Application US/08185851A  
Patent No. 5489430  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Ohkawa, Setsuko  
APPLICANT: Fujisawa, Ayumi  
APPLICANT: Iitani, Yoshikazu  
APPLICANT: Aoyama, Shigemitsu  
TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene  
TITLE OF INVENTION: Thereof and Recombinant Vectors Containing the Gene As  
TITLE OF INVENTION: Well As Vaccines Utilizing the Same  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Armstrong, Westerman, Hattori, Mcleland &  
ADDRESS: Naughton  
STREET: 1725 K Street, N.W., Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0  
SOFTWARE: ASCII from Word Perfect version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,851A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Theresa M. Stevens-Smith  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: PO-8-A930918  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-887-0357  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-185-851A-4

Query Match 69.7%; Score 1619; DB 1; Length 368;  
Best Local Similarity 95.5%; Pred. No. 5.8e-106;  
Matches 322; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 64 CMSITKQDANPNNGQTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSAEATVN 123  
DB 27 CMSITKQDANPNNGQTQLEAARMELTDLINAKAMTLASLODYAKIEASLSAYSAEATVN 86  
QY 124 NNLAATLEQLMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLA 183  
DB 87 NNLAATLEQLMAKTNLESAINQANTDKTTPDNEHPNVEAYKAKTTLLEORATNLEGLA 146  
QY 184 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEOKTN 243  
DB 147 STAYNOIRNNLVLDLYNKAASLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEOKTN 206  
QY 244 ADALNSFIKKVYQNNQSFVGTFTNANVQPSNYSFVAFSADVTVPVYKVARRTVWNGDE 303  
DB 207 ADALNSFIKKVYQNNQSFVGTFTNANVQPSNYSFVAFSADVTVPVYKVARRTVWNGDE 266  
QY 304 PSSRLIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKADANNVGLQ 363  
DB 267 PSSRLIANTNSITDVSWIYSLAGTNTKYQFSFSNYGPGSTGYLYFPYKLVKADANNVGLQ 326

QY 364 YKLNNGNVQVEFATSTANTTANPTAVDEIKYAK 400  
Db 327 YKLNNGNVQVEFATSTANTTANPTAQMLKILK 363

## RESULT 5

US-08-525-742-2  
Sequence 2, Application US/08525742  
Patent No. 5871742  
GENERAL INFORMATION:  
APPLICANT: Saito, Shuji  
APPLICANT: Onkawa, Setsuko  
APPLICANT: Saeki, Sakiko  
APPLICANT: Ohsawa, Ikuroh  
APPLICANT: Funato, Hiroko  
APPLICANT: Iitani, Yoshikazu  
APPLICANT: Aoyama, Shigemasa  
APPLICANT: Takahashi, Kiyoohito  
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
ADDRESS: NAUGHTON  
STREET: 1725 K Street, Suite 1000  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,742  
FILING DATE: 25-SEP-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-074139  
FILING DATE: 31-MAR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 05-245625  
FILING DATE: 30-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP94/00541  
FILING DATE: 31-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcleland, Le-Nhung  
REGISTRATION NUMBER: 31,541  
REFERENCE/DOCKET NUMBER: 950811  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-659-2930  
TELEFAX: 202-8870357  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-525-742-2

Query Match 34.6%; Score 803; DB 2; Length 661;  
Best Local Similarity 41.4%; Pred. No. 2e-48;  
Matches 187; Conservative 64; Mismatches 129; Indels 72; Gaps 13;

QY 51 PTVGCTVITLRFSGCSITTKDANPNNGQQLQLEAARELTD-----LINAQAMTLASIQ 103  
Db 35 PPSGG-----NMNGGNTNTPSDGGQGMNMAAAKELADAKAALTTLLINGETANLASYE 84  
QY 104 DYAKIEASISAVSEAEATVNNNLNATLEQLKMAKTNLSAIAQANTDKTFPNEHNLVE 163

Db 85 DYAKIESELISAYTATAVAKATGATLNEVNEAKTTLDAAIKKAASAKNPFDAQHSGLVE 144  
QY 164 AYKALKTTLEQATNEGLSSSTAYNQIRNMLVDLYNKAASSLITKTLDPGLNGTLLDSNEI 223  
Db 145 AYNMLKEKLEKEKNTLDSLANENYAAIRTNLNSLYEKANTIVATLDPAT-GNPEWMSV 203  
QY 224 TTVNRNINNTLSTINECKTNADLNSGFIKKVIONNESFGFTNTANVOPSNYSFAVS 283  
Db 204 TOANQDITNATSTRILAKQNDNLANSFIKOSLVKNMLRVDVANNQ-OPANYSFVGS 262  
QY 284 ADVTPVNYKYARRTVMNGD-----PSSRLANTNSITDVSWIYSLAGNTTKYQFSPSN 337  
Db 263 VNVDTPMNMFQKRVKMSSEMTPLATTPAEDATQQAASLTDVSWIYSLNGAEAKYTLSPRY 322  
QY 338 YG-PSGTLYFPYLYVAAADANNTGLQYKLNNGVQVEF-----ATSSAN--NTA 387  
Db 323 FGAERTAVLYFPYLYVTSQD--NVGLQYKLNNGGTQKINFVQTPASGSSDVAANEESTMA 380  
QY 388 NP-----TPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNMNKVAPMIGNIYLS- 438  
Db 381 SPAMQSPVVDITKIAKVALSNLKNSTIERSVPTG-----KAEMIGNIYLTSSNSE 435  
QY 439 -----NENN-----ADKIPGY 449  
Db 436 VNKKIYDDLFGNSFNNENPTAVTVDLKGY 467

## RESULT 6

US-07-803-633A-13  
Sequence 13, Application US/07803633A  
Patent No. 5369025  
GENERAL INFORMATION:  
APPLICANT: NAZERIAN, Keyvan  
APPLICANT: LEE, Lucy F.  
APPLICANT: YANAGIDA, No. 5369025oru  
APPLICANT: OGAWA, Ryohel  
APPLICANT: LI, Yi  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VACCINE FOR  
TITLE OF INVENTION: PROTECTION AGAINST MAREK'S DISEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 No. 5369025ch Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/803,633A  
FILING DATE: 19911210  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerold M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1644-103P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 865 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-803-633A-13

1. **Introduction**  
 2. **Method**  
 3. **Results**  
 4. **Discussion**  
 5. **Conclusion**  
 6. **References**  
 7. **Appendix**  
 8. **Tables**  
 9. **Figures**  
 10. **Supplementary Materials**  
 11. **Notes**  
 12. **References**  
 13. **Appendix**  
 14. **Tables**  
 15. **Figures**  
 16. **Supplementary Materials**  
 17. **Notes**  
 18. **References**  
 19. **Appendix**  
 20. **Tables**  
 21. **Figures**  
 22. **Supplementary Materials**  
 23. **Notes**  
 24. **References**  
 25. **Appendix**  
 26. **Tables**  
 27. **Figures**  
 28. **Supplementary Materials**  
 29. **Notes**  
 30. **References**  
 31. **Appendix**  
 32. **Tables**  
 33. **Figures**  
 34. **Supplementary Materials**  
 35. **Notes**  
 36. **References**  
 37. **Appendix**  
 38. **Tables**  
 39. **Figures**  
 40. **Supplementary Materials**  
 41. **Notes**  
 42. **References**  
 43. **Appendix**  
 44. **Tables**  
 45. **Figures**  
 46. **Supplementary Materials**  
 47. **Notes**  
 48. **References**  
 49. **Appendix**  
 50. **Tables**  
 51. **Figures**  
 52. **Supplementary Materials**  
 53. **Notes**  
 54. **References**  
 55. **Appendix**  
 56. **Tables**  
 57. **Figures**  
 58. **Supplementary Materials**  
 59. **Notes**  
 60. **References**  
 61. **Appendix**  
 62. **Tables**  
 63. **Figures**  
 64. **Supplementary Materials**  
 65. **Notes**  
 66. **References**  
 67. **Appendix**  
 68. **Tables**  
 69. **Figures**  
 70. **Supplementary Materials**  
 71. **Notes**  
 72. **References**  
 73. **Appendix**  
 74. **Tables**  
 75. **Figures**  
 76. **Supplementary Materials**  
 77. **Notes**  
 78. **References**  
 79. **Appendix**  
 80. **Tables**  
 81. **Figures**  
 82. **Supplementary Materials**  
 83. **Notes**  
 84. **References**  
 85. **Appendix**  
 86. **Tables**  
 87. **Figures**  
 88. **Supplementary Materials**  
 89. **Notes**  
 90. **References**  
 91. **Appendix**  
 92. **Tables**  
 93. **Figures**  
 94. **Supplementary Materials**  
 95. **Notes**  
 96. **References**  
 97. **Appendix**  
 98. **Tables**  
 99. **Figures**  
 100. **Supplementary Materials**  
 101. **Notes**  
 102. **References**  
 103. **Appendix**  
 104. **Tables**  
 105. **Figures**  
 106. **Supplementary Materials**  
 107. **Notes**  
 108. **References**  
 109. **Appendix**  
 110. **Tables**  
 111. **Figures**  
 112. **Supplementary Materials**  
 113. **Notes**  
 114. **References**  
 115. **Appendix**  
 116. **Tables**  
 117. **Figures**  
 118. **Supplementary Materials**  
 119. **Notes**  
 120. **References**  
 121. **Appendix**  
 122. **Tables**  
 123. **Figures**  
 124. **Supplementary Materials**  
 125. **Notes**  
 126. **References**  
 127. **Appendix**  
 128. **Tables**  
 129. **Figures**  
 130. **Supplementary Materials**  
 131. **Notes**  
 132. **References**  
 133. **Appendix**  
 134. **Tables**  
 135. **Figures**  
 136. **Supplementary Materials**  
 137. **Notes**  
 138. **References**  
 139. **Appendix**  
 140. **Tables**  
 141. **Figures**  
 142. **Supplementary Materials**  
 143. **Notes**  
 144. **References**  
 145. **Appendix**  
 146. **Tables**  
 147. **Figures**  
 148. **Supplementary Materials**  
 149. **Notes**  
 150. **References**  
 151. **Appendix**  
 152. **Tables**  
 153. **Figures**  
 154. **Supplementary Materials**  
 155. **Notes**  
 156. **References**  
 157. **Appendix**  
 158. **Tables**  
 159. **Figures**  
 160. **Supplementary Materials**  
 161. **Notes**  
 162. **References**  
 163. **Appendix**  
 164. **Tables**  
 165. **Figures**  
 166. **Supplementary Materials**  
 167. **Notes**  
 168. **References**  
 169. **Appendix**  
 170. **Tables**  
 171. **Figures**  
 172. **Supplementary Materials**  
 173. **Notes**  
 174. **References**  
 175. **Appendix**  
 176. **Tables**  
 177. **Figures**  
 178. **Supplementary Materials**  
 179. **Notes**  
 180. **References**  
 181. **Appendix**  
 182. **Tables**  
 183. **Figures**  
 184. **Supplementary Materials**  
 185. **Notes**  
 186. **References**  
 187. **Appendix**  
 188. **Tables**  
 189. **Figures**  
 190. **Supplementary Materials**  
 191. **Notes**  
 192. **References**  
 193. **Appendix**  
 194. **Tables**  
 195. **Figures**  
 196. **Supplementary Materials**  
 197. **Notes**  
 198. **References**  
 199. **Appendix**  
 200. **Tables**  
 201. **Figures**  
 202. **Supplementary Materials**  
 203. **Notes**  
 204. **References**  
 205. **Appendix**  
 206. **Tables**  
 207. **Figures**  
 208. **Supplementary Materials**  
 209. **Notes**  
 210. **References**  
 211. **Appendix**  
 212. **Tables**  
 213. **Figures**  
 214. **Supplementary Materials**  
 215. **Notes**  
 216. **References**  
 217. **Appendix**  
 218. **Tables**  
 219. **Figures**  
 220. **Supplementary Materials**  
 221. **Notes**  
 222. **References**  
 223. **Appendix**  
 224. **Tables**  
 225. **Figures**  
 226. **Supplementary Materials**  
 227. **Notes**  
 228. **References**  
 229. **Appendix**  
 230. **Tables**  
 231. **Figures**  
 232. **Supplementary Materials**  
 233. **Notes**  
 234. **References**  
 235. **Appendix**  
 236. **Tables**  
 237. **Figures**  
 238. **Supplementary Materials**  
 239. **Notes**  
 240. **References**  
 241. **Appendix**  
 242. **Tables**  
 243. **Figures**  
 244. **Supplementary Materials**  
 245. **Notes**  
 246. **References**  
 247. **Appendix**  
 248. **Tables**  
 249. **Figures**  
 250. **Supplementary Materials**  
 251. **Notes**  
 252. **References**  
 253

RESULT 9  
US-09-421-868-2  
; Sequence 2, Application US/09421868  
; Patent No. 6177084  
; GENERAL INFORMATION:  
; APPLICANT: Foster, Timothy J.

APPLICANT: McDevitt, Damien L.  
TITLE OF INVENTION: The S. aureus Fibrinogen Binding Protein Gene  
FILE REFERENCE: 05344.105011  
CURRENT APPLICATION NUMBER: US/09/421,868  
CURRENT FILING DATE: 1999-10-19  
PRIOR APPLICATION NUMBER: 08/293,728  
PRIOR FILING DATE: 1994-08-22  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 2  
LENGTH: 933  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-421-868-2

Query Match 7.6%; Score 176.5; DB 3; Length 933;  
Best Local Similarity 22.7%; Pred. No. 0.00026;  
Matches 112; Conservative 66; Mismatches 184; Indels 131; Gaps 23;

QY 2 HYFRNCIFFLVIL-----YGTNS-----PSTONVTSREVSVQLSSESTFYLCP 52  
DB 9 HAIRKKSIGVASVLGTIGFGLSSKEADASENSVTSQDSASNSKSDSSVSAAPKT 68  
QY 53 VGSIVIRLEFGGMSITKKDAPNNGQDLP--EAMMELTDLINAKAMTLASLDYAKIEA 110  
DB 69 DDTNV-----SDTKSSNTNNGETSVAQNPAGQETTSSTNATL----- 108  
QY 111 SLSSASVSEAEVYNNNLATLEQLKAKTNLESAINQANTDKTTPNEHNPVLEAYKALKT 170  
DB 109 ESTPTGATTTTITQANTPATTOSSNTVAEELVNO--TSNETTFND--TNTVSSVNS--- 162  
QY 171 TLEORATNLEGLSR-----AYNQRNNLVVLNKK-----ASSLIT 206  
DB 163 --PQSTINAEVSTIQDTSTEARPSNBSAPQSTASNDVNOAVNTSAPMRAFSIAA 220  
QY 207 KTLDELNGTLLDSNEITTVNENINNTLSTINEQ---KTN--ADALNSFIK---KVIO 257  
DB 221 VAADAPAGTDI--TNQLTNVTGIDSGTTPVPHQGYVLNAGFSPVSAVAGDFPKITV 279  
QY 258 NNEQSFVGTFTNANVQPSNYSFVAASADYTPVNYKARRTVNGCEPSSRIANTNSITD 317  
DB 280 PKEILNNGVTSSTAKPP-----IMAGQDQ---VLANGVIDSD 312  
QY 318 VSMWISLNG--TNTKYQSFSPNNGSPSTGYLYFPYKLVKADANNVGLQYKLNNGVQOYEF 376  
DB 313 GNVITTFIDYVNTKODVKATLTMPA--YI-----DPERV-----KKTGNV----- 350  
QY 377 ATSTSANNTANPFAVDEIKVAKI-----VLSGLRFGONTIELSVPTGEGNNKVAP 429  
DB 351 TLATGISTTANKTVLVDEYKGYKFPNLSIKGTIDQIDKTNNTYRQTIYVNSGDNVIA 410  
QY 430 MI-GNIVYSSNEN 441  
DB 411 VLTGMLKENTDSN 423

RESULT 10  
US-08-220-151-4  
Sequence 4, Application US/08220151  
Patent No. 5529780  
GENERAL INFORMATION:  
APPLICANT: Paololetti, Enzo  
APPLICANT: Limbach, Keith J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESS: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/220,151  
FILING DATE: 30-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2540  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
TELEX: 425066 CURTMS  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1041 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-220-151-4

Query Match 7.5%; Score 174.5; DB 1; Length 1041;  
Best Local Similarity 19.8%; Pred. No. 0.00041;  
Matches 118; Conservative 70; Mismatches 188; Indels 221; Gaps 24;

QY 3 YFRNCIF-----FLIVLY----- 17  
DB 22 YFRQCFPPSLGLAAGSRHNGSSGITLARVSTFWLFLVGPDPVGGSGSTSEQ 81  
QY 18 -----GTNSPSTONVTSREVSVQLSSEB---STFYLCPVGSIVIRLE- 61  
DB 82 PRRTVATPEVGCTPPKPTPTDPTDMSDMREALRAQIEANGSTYMCPPGSGTVRLRP 141  
QY 62 -----FGCMSTKKDANP-----NNGTQ 80  
DB 142 PRACPDYKLGKPTFEGIAVIFKENIAPYKFNAYIYNNIITTWSSGSSVAVTTRTYDR 201  
QY 81 LEARMELTDLINAKAMTLASLDYAKIEALSAYSSEAEFTVNNNLATLEQLKAKTNL 140  
DB 202 VPVAVQETIDILDRKML-----SKADYVRNNYQFTAPDRBDEPREL 244  
QY 141 ESAINQANTDKTTPDNEHNPVLEAYKALKTLEQRATNLEGLSSTAYNOIRNNLVDLNKK 200  
DB 245 PLKPSKEVTPQSR--GWH-----TYKFAVYVYKDVIVSTAMAGSSYTQITNRVADRVP 297  
QY 201 ASSLITKTLDELNGTLLDSNEITTVNENINNTLSTINEQNTND-----ALNSFIKV 255  
DB 298 PVSEITDTIDKFG-----KCSSKATVYRN--NHKVEARFEDKNPQDMLIAKSNVSGSKA 351  
QY 256 IONNEQSFV-----GFTNANVQP---SNYSFVAFSADVTTPVNY----- 291  
DB 352 WHTNREYTKGAAGFHHSGTSVNCIYBEVDARSVIPDPSAISTGVYIHMSPPFGLRDG 411  
QY 292 KYARRTYWNGD-----PSSRIANTNSITDYSW-----IYS 323  
DB 412 AHVHTSYSSDRFQOIGGYPIDDTRLQAGAPSRNPLETPHVT--VAMNMTPKGRVCT 470  
QY 324 LAG-----TNTKYQSFSPNNGSPSTGYLYFPYKLVKADANNVGLQYKLNNGVQOYEF 376  
DB 471 LAKWREIDEMLRDEYQGSYR-----FTVTTIATPFISSNTS--QFEIV--RIRLGDC 517  
QY 377 ATSTSANNTANPFAVDEIKVAKIIVLSGLRFGONTIELSVPTGEGNNKVAPIGN 433  
DB 518 ATKEAAL-----AIDRIYKSKYKSTHIQTG--TLETYIARG--GFLIARPMISN 563

RESULT 11

US-08-413-118-4  
Sequence 4, Application US/08413118  
Patent No. 568920  
GENERAL INFORMATION:  
APPLICANT: PROLETTI, ENZO  
APPLICANT: LIMBACH, KEITH J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE, 25TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,118  
FILING DATE: 29-MAR-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/220,151  
FILING DATE: 30-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: FROMMER, WILLIAM S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1041 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-413-118-4

Query Match 7.5%; Score 174.5; DB 1; Length 1041;  
Best Local Similarity 19.8%; Pred. No. 0.00041;  
Matches 118; Conservative 70; Mismatches 188; Indels 221; Gaps 24;

3 YFRNCIF-----FLVILY----- 17  
22 YFRQRCFPPLSLGIAATGSRHNGSSGLTRLARVYFVILVFLVGRPPVGGSGSTSEQ 81  
18 -----GYNSSPSTQNTVTSREVVSVQLSEB---STFYLCPPVGSIVIRLE- 61  
82 PRRVATPEVGSTPKPTTDPDMDREBALRASQIEANGPSTFCWCPGSGSTVRLRP 141  
62 -----FGCMSITKKDANP-----NNGQTQ 80  
142 PRACPDYKLGKNGFTGIAVFIEKEMIAPYKFKANIYKNIIMTTWSSGSSYAVATTIRYDR 201  
81 LEAAMELTDLINAKMTLASLQDYAKIEASLSAYSEAEYNNNNLNTLEQLMAKNTNL 140  
202 VPVKQETITDLIDRGMCLE-----SKADYRNNYQGTAPRDEDPREL 244  
141 ESAINQANTDKTTTFNEHPNLVEAYKALKITLQGRATNLEGLSTAYNQIRNNLVLDLYNK 200  
245 PLKPSKFTPPQSR---GMH-----TYKFKATVYKDVIVSTAMAGSSYQITFRYADRVPI 297  
201 ASLITKTLPLNGTGLDSNITVNNINNTLSTINEOKTNAD-----ALSNFIRKV 255  
298 PVSEITDITDKFG-----KCSSKATVVRN-NHKVEAFNEKDPQDMLASKYNSVGSKA 351

256 IONNEQSFV-----GTFNANVOP-----SNYSFVARSADVTTPVNY----- 291  
352 WHTNNEYTKLGAGFPHSGTSVNCIYEBVDARSVPDPSAISTGVYIHMSPFEGLRDG 411  
292 KYARRTYWNGDE-----PSSRLIANTNSTITDSM-----IYS 323  
412 AHVHTSYSSDRFOQIEGYPIDLTRLQLGAPVSRNPLEPHVT-VAMNWTPKGRVCT 470  
324 LAG-----TNTYQPSFSNYPGSTGVLPVYKLVFAADANVGLQYKLNNGVQVEF 376  
471 LAKREIDEMLRDEYQGSYR-----FTVKTISATPISNTS-OFEIN-RIRLQDC 517  
377 ATSTANNTTANPPPAVDKIKVAKIVLSGLRFGQNTIELSVPTSGNNKVKVPMIGN 433  
518 ATKRAAE-----AIDRIYKSKYSKTHIQTG--TLETYLARG-GFLIARPMISM 563

RESULT 12  
US-08-473-446-4  
Sequence 4, Application US/08473446  
Patent No. 6017542  
GENERAL INFORMATION:  
APPLICANT: PROLETTI, ENZO  
APPLICANT: LIMBACH, KEITH J.  
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF  
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR  
NUMBER OF SEQUENCES: 128  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.  
STREET: 530 FIFTH AVENUE, 25TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,446  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/413,118  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FROMMER, WILLIAM S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2670  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1041 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-473-446-4

Query Match 7.5%; Score 174.5; DB 3; Length 1041;  
Best Local Similarity 19.8%; Pred. No. 0.00041;  
Matches 118; Conservative 70; Mismatches 188; Indels 221; Gaps 24;

3 YFRNCIF-----FLVILY----- 17  
22 YFRQRCFPPLSLGIAATGSRHNGSSGLTRLARVYFVILVFLVGRPPVGGSGSTSEQ 81  
18 -----GYNSSPSTQNTVTSREVVSVQLSEB---STFYLCPPVGSIVIRLE- 61

Db 82 PRRTVATPEVGTGTPKPTTDDPTDMSDMREALRASQIEANGSPSTFMCPPSSGSTVVRLEP 141  
Qy 62 -----FGCMSTTKQDANP-----NNGQO 80  
Db 142 PRACPDYKLGKNTGTEGIAVIFKENIAPYKFKANIYYKNIIMTWGSSSAYATTNNYTR 201  
Qy 81 LEAARMELTDLINAKAMTTLASLODYAKIEASLSAYSBEATVNNINATLEOLKMAKATL 140  
Db 202 VPKVQVEITDILIRKMCLE-----SKADYANNNTQFTAPDRDEDPREL 244  
Qy 141 ESAINQANTDKTTFDNEHNPVAYKALKTLEQORATNLEGLSTAYNOIRNNLVLYNK 200  
Db 245 PLKPSKFTPQSR--GMH-----TYKFKATVYKQDIYVSTAMAGSSYTOITNRADRVPI 297  
Qy 201 ASSLIKTIDPLNGGLDLSNETTNNRNINNTLSTINECKTAD-----ALSNSTKCV 255  
Db 298 PVSEITDITDKFG-----KCSSKATVYRN-NHKVEAFNDKQFQDPLIASKXNSGSKA 351  
Qy 256 IONNEOSFV-----GTFNNANVP-----SNYSFVAFSADVTPVNY-----291  
Db 352 WHTTNTYTKIGAGFHSGTSVNCIVEVDARSVPYPSFALSTGDVTHMSPPGLRQD 411  
Qy 292 KXARTVMNGD-----PSSRIANTNSITDVS-----IYS 323  
Db 412 AHVEHTSYSGSDPFOQIEGYPIDLDTROLGAPVSHNPLETPHVT-VANNMTPKCGRVCT 470  
Qy 324 LAG-----TNTKQFSFSNYPSTGYLYFPYKLVKADANNVGLQYTLANNVQOUEF 376  
Db 471 LAKWEIDEMLDEYVGSYR-----FTVKITISAFISNTS-QFEIN--RRLGDC 517  
Qy 377 ATSTSANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAPMIGN 433  
Db 518 ATKEAAE-----AIDRIYKSKYSKTHIQG--TLERYLARG-GLIARFPMISN 563

RESULT 13  
US-08-836-687B-37  
; Sequence 37, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John  
; APPLICANT: Dodd, Helen Mair  
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN  
; FILE REFERENCE: 20747/70  
; CURRENT APPLICATION NUMBER: US/08/836,687B  
; CURRENT FILING DATE: 1995-11-20  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 37  
; LENGTH: 682  
; TYPE: PR1  
; ORGANISM: Lactococcus sp.  
US-08-836-687B-37

Query Match 6.5%; Score 150.5; DB 4; Length 682;  
Best Local Similarity 21.0%; Pred. No. 0.01;  
Matches 75; Conservative 60; Mismatches 131; Indels 91; Gaps 13;

Qy 120 ETVNNNNLAAATLEOLMAKTNLESAINQANTDKTTFDNEHNPVAYKALKTLEORATNL 179  
Db 38 ELINNSNNALISSTGISTDSTINLGAOSPAVAKSTRTB--LDVTGAKATLLQISAQOK 93  
Qy 180 EGLSTAVNOIRNNLVLYNKASSLITKLDPLNGGLDLSNETTNNRNINNTLSTINE 239  
Db 94 EAKVSLQETQVSS--FSKRDSTVTKAEAVPGDELLEQSEVAVVSSIQK-KILDN 148  
Qy 240 OKTNADALSNSFIKVIQNNNEOSFVGTNNANVPNSYFVAFSADVTPVNYKXARTV- 298  
Db 149 KKRANFVTSFPLIKKPSNSKDSAGVIDNS-----ASPLSYRRAKEVVS 193  
Qy 299 -----NNGDEPSSRIIANTNSITDVSWISLACTNFKYQPSNSNPGS 341  
Db 194 LRQPLAKQKVEAOPILISNSSEKKASVYTNSHDEWDYQW-----DMKY--VTNNGES 243

Qy 342 TGYLYFPYKLVKADANNVGLQYKLNNGNVQVEATSTAN-----NT 385  
Db 244 YA-LYQPSKKI-----SVGI-----IDSGIMEEHPLDLSLGNYPKVLVPGKGFDEBDE 293  
Qy 386 TANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAPMIG-NITLSSNEN 441  
Db 294 TGNPSDIDV-----KMGHTEVAQOITANGNIIQVAPQITVNIYRVGEN 338

RESULT 14  
US-09-541-782-6  
; Sequence 6, Application US/09541782  
; Patent No. 6284480  
; GENERAL INFORMATION:  
; APPLICANT: Nislow, Corey  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Beraud, Christophe  
; TITLE OF INVENTION: Antifungal Assay  
; FILE REFERENCE: 1015  
; CURRENT APPLICATION NUMBER: US/09/541,782  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1073  
; TYPE: PR1  
; ORGANISM: S.pombe  
US-09-541-782-6

Query Match 6.3%; Score 147.5; DB 3; Length 1073;  
Best Local Similarity 20.0%; Pred. No. 0.034;  
Matches 109; Conservative 91; Mismatches 203; Indels 141; Gaps 25;

Qy 17 YGNSSPSTQNT-----SREVSQV-----QISEESTFYLCPPPVGSTVIR 59  
Db 542 YETNEAKITVATDLSQYRESKEYIASLYEKLDRTERRNNKENNNF-----N 590  
Qy 60 LERCGMSITK-----KDNPNNGO--TOLEARMELTDLINAKA-MTLASL-----ODYAKIE 109  
Db 591 LKFNILTMLSFHSFGFTDETNGYFTLLDNPASMEELNTHSNQLISMKITEHPQSLD 650  
Qy 110 ASLSAYSEAEVTNNINATLEOLMAKTNLESAINQANTDKTTFDNEHNP-----LVEA 164  
Db 651 EALOSASSQCAVPSISDLYSELKSKSLDLALHSLODISMSQKLGNGISSEIETL 710  
Qy 165 YKALKTT--LEORATNLEGLSTAVNOIRNNLVLYNKASSLITKLDPLNGGLDLSN 221  
Db 711 QOKWESYRQVLQELRSLVNLQHTHEESQKELMVGANDIDALVKTCTSLNDADITLSD 770  
Qy 222 EIT-----TNNRNINNTLSTINEQ-KTNADAL-----SNSFIKK--VI 256  
Db 771 YISDQSKSPSKQODLIANTIGKIVSNFLQONESLYTKADILSHLNDTNSNRKANEIM 830  
Qy 257 QNNEOSFVGTNNANVPNSYFVAFSADVTPVNYKXARTVNG-----DEPSSRIANT 312  
Db 831 NNNSEEL--RNA-----ASQAEIVGANKKEIOCTVENGSLDLSKKAHSNS 877  
Qy 313 NSITDVSWISLA-----GNTKYO-----PFSNYSYGSTGYLYFPYKLVKA 354  
Db 878 RSWYD-----HCLALAESQOGVNLLEVOTLDRLLQKVEHSEDNTKEHQQL--LDLLES 930  
Qy 355 ADANNVGL--QYKLNNGNVQVE-----FATSTANNT-----ANPTPAVDEIKV 398  
Db 931 LVGNNDNLISITKPHTELQITDHYLKGTTSIANHTELLGLGDSLCLETTIEDTSL 990  
Qy 399 AKIVLSGLRFGONTI-----ELSVPTGEGNNKVAPMIGNIYLSNENNAKIPG 448  
Db 991 VKLETTGDTPSKRELPAITPSWTRDSSLIKETTNMLNDSQKFRFRTYTSNQNNEPDV-- 1048  
Qy 449 YRRP 452  
Db 1049 YDKP 1052

RESULT 15  
IIS-09-723

US-09-723-820-6

; Sequence 6, Application US/09723820

Patent No. 6468760

GENERAL INFORMATION:

APPLICANT: Nislow, Corey

APPLICANT: Sakowicz, Roman

APPLICANT: Beraud, Christophe

; TITLE OF INVENTION: Antifungal Assay

FILE REFERENCE: 1015

CURRENT APPLICATION NUMBER: US/09/723,820

CURRENT FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: 09/541,782

PRIOR FILING DATE: 2000-04-03

NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1073

TYPE: PRT

ORGANISM: *S. pombe*

US-09-723-820-6

Query Match	Score	DB	Length
6.3%	147.5	DB 4	1073

Best Local Similarity 20.0%; Pred. No. 0.034;

Matches 109; Conservative 91; Mismatches 203; Indels 141; Gaps 25;

```

0Y 17 YGINSPESTQANT-----SREAVSV-----QISEESTFLCPPEVASTYIR 59
Db 542 YEINAEKITYATVDLSOYYRESKEIYALYELDTERKNKENENNFM-----N 590
0Y 60 LEFGMSITK-----KDANPNNGO--TQLEAARMELTDLINAKA--MTLAL-----ODYAKIE 109
Db 591 LKFNULTWLRSPHSGFTDETNGYFLLNDPNAMSMELLNTHSNOLLIMTKTEHPQSLD 650
0Y 110 ASLSAYSBEATVNNNNNATLEQLKMAKTULESAINQANTDKTTPDNHPN-----LYEA 164
Db 651 EALOSARSCCAVPNSLDLIVSELKDSKNSLDDALEHSILODISMSOKLGNIGISELIEL 710
0Y 165 YKALKTT---LEQRATNLEGSTAVNQIRNNVLVLYNKASLSIKTILDPNGGTLUDSN 221
Db 711 QDKMKESIYROLVQELRSLNLOHTHEESQKELMTGVNRVDIDLVTXCTTSLNDADITLSD 770
0Y 222 EIT-----TVNRNINNTLSTINEQ--KTNDAL-----SNSEIKK-----VI 256
Db 771 YISDPKSKFESKQODLIANIGKIYVNFQEOENESLYTADILHSHLNTNINSIRAKEM 830
0Y 257 QNNEOSFVGTFTTANVOPNSNVSFAFSADVTVPVNYKVARRYWNG-----DEBSSHILANT 312
Db 831 NNRSEEPF--RNA-----ASQAEIVGANKERIOKTVENSOQLDCKSKALIHNS 877
0Y 313 NSIDVSMYISLA-----GNTTKYQ-----FSFENYSPSPGILYFPKLYKA 354
Db 878 RSMYD---HCLALAEQKQGVNLEVOQLDRLLOKVEKHESEDNTRYEKHQOL---LDLES 930
0Y 355 ADANNVGL--QYKLANGVNOVE---PATSTSANNT-----ANPTPAVDKIV 398
Db 931 LVGNNDNLIDSKTTPHTELOKITDHVLKGTISLHNTHEBLGQGESICLNLETTIEDTSL 990
0Y 399 AKIVLSGLRFGONTI-----ELGVPTEGNNKVAAPMIGNIYLSNEMNADKIPG 448
Db 991 VKLETTGDTPSKRELPATPSPWTRDSSLKETNNLMLDSQKKFVRERTYSSNQTNPEPV-- 1048
0Y 449 YRRP 452
Db 1049 YDKP 1052

```

Search completed: August 14, 2003, 10:24:53  
Job time : 32 secs

**THIS PAGE BLANK (USPTO)**





Qy	661	CGAATAAATTAGTGGATCTATACATAAAGCTGAGTTAATAACTAAACACTAGAT	720
Db	661	CGAATAAATTAGTGGATCTATACATAATCTGTAGTTAAATACTAAACACTGAT	720
Qy	721	CCACTAAATGGGGGAACGCTTTAGATCTPATGAGATTACTACAGTTATCGGAATATT	780
Db	721	CCACTAAATGGGGGAATGCTTTAGATTCTAATGAGATTACTACAGTTAATCGGAATATT	780
Qy	781	AATAATAGCTTATCAACTATTATATGACAACAAAAGCTAATGCTGATGATTCATTAATAGT	840
Db	781	AATAAATACCTTTCAACTATTATATGACAACAAAAGCTAATGCTGATGATTCATTAATAGT	840
Qy	841	TTTATTAATAAAGATGATTCAAAATTAATGAACAAGTTTGTAGGAGCTTTTACAAACGCT	900
Db	841	TTTATTAATAAAGATGATTCAAAATTAATGAACAAGTTTGTAGGAGCTTTTACAAACGCT	900
Qy	901	AATGTTCACCTTCAAACTACAGTTTGTGCTTTAGTGCATGTGAACCCCGTCAT	960
Db	901	AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGCATGTGAACCCCGTCAT	960
Qy	961	TATTAATATGCAAGAAGACCGTTTGGATGATGTGAACCTTCAGTAGAATTCTTGGCA	1020
Db	961	TATTAATATGCAAGAAGACCGTTTGGATGATGTGAACCTTCAGTAGAATTCTTGGCA	1020

[illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

—

[illegible][illegible]

1201 GTTCAACAAAGTTGAGTTTGCACCTTCAACAAAGTGCACATATATACCTACACCTAATCCAACT 125

Db 1201 GTTCAACAGTTGAGTTTGCCACTTCAACTAGTCAATAATACTACAGCTAATCCACT 1266

1261 CCAGCAGTTGATGAGATTAAGTTGCTAAATCGTTTATCAGGTT 1306

Db 1261 CCAGCAGTTGATGAGATTAAGTTGCTAAATCGTTTATCAGGTT 1306

AAQ77854

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	

XX 2E-MXD 2002 (11/23/02)

DT 23-JUN-1995 (first entry)  
XY

DE Mycoplasma gallisepticum 40kd antigen coding sequence.  
XX

recombinant avipox virus; live vaccine; mycoplasma 40kd antigen; TTM-1: ds.

XX  
OS *Mycoplasma gallisepticum*.

aa	Key	Location/Qualifiers
FH		

```

C00      202..1300
FT      /*tag= a

```

FT /note= "ORF includes 2 NNN codons"

PN WO9423019-A1.

PD 13-OCT-1994.

PF 31-MAR-1994; 94WO-JP00541.  
VY

PR 31-MAR-1993; 93JP-0074139.  
 PR 30-SEP-1993; 93JP-0245625.  
 PA (JAPG) NIPPON ZEON KK.  
 PA (SHIO) SHIONOGI & CO LTD.  
 PI Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;  
 PI Saeki S, Saitos, Takahashi K;  
 DR WPI: 1994-333181/41.  
 DR P-RSDB; AAR63227.  
 XX Recombinant avipox virus combining DNA encoding a polypeptide -  
 PT exhibiting antigenicity of mycoplasma, useful for the production  
 PT of a live vaccine  
 XX  
 PS Claim 4; Page 71-74; 123pp; Japanese.  
 CC The plasmid pUTM-1P contains a sequence (the TTM-1 gene) coding  
 CC for the 40kD antigen of Mycoplasma gallisepticum under the control  
 CC of a synthetic promoter. A 1300 bp restriction fragment containing  
 CC the promoter-ORF sequence was excised and was used in the  
 CC construction of plasmid pM7929-R2. This in turn was involved in the  
 CC construction of a recombinant avipox virus vector comprising the  
 CC TTM-1 gene, DNA encoding the signal membrane anchor peptide from  
 CC Newcastle Disease Virus haemagglutinin neuraminidase and FPV  
 CC sequences. The recombinant avipox virus is useful as a live vaccine  
 CC to protect against infection by Mycoplasma gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct FN field.)  
 CC  
 XX  
 SQ Sequence 1387 BP; 531 A; 220 C; 203 G; 427 T; 6 other;  
 Query Match 96.8%; Score 1264; DB 15; Length 1387;  
 Best Local Similarity 98.3%; Pred. No. 5, 4e-196;  
 Matches 1284; Conservative 0; Mismatches 21; Indels 1; Gaps 1;  
 QY 1 AAAAAATCATGATGTTTATCTGATATCTTTGCTTAAAAAACAACAAATCTTCTAACAA 60  
 DB 1 AAAAAATCATGATGTTTATCTGATATCTTTGCTTAAAAAACAACAAATCTTCTAACAA 60  
 QY 61 AATCTTAAATTAATAGCGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120  
 DB 61 AATCTTAAATTAATAGCGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120  
 QY 121 AACCAAAATCTCTAGTAAAGCGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 180  
 DB 121 AACCAAAATCTCTAGTAAAGCGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 180  
 QY 181 ATAAATATATCTTAAATTTCTATGATGAAGAAAGAAATCATCTTAAAGACTATTAAGTTG 240  
 DB 181 ATAAATATATCTTAAATTTCTATGATGAAGAAAGAAATCATCTTAAAGACTATTAAGTTG 240  
 QY 241 TTAGGTACAAATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 300  
 DB 241 TTAGGTACAAATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 300  
 QY 301 GATGCAAAACCAAAATTAATGCGCAAAACCAATTAAGAGAGCGGCAATGAGTTAACAGAT 360  
 DB 301 GATGCAAAACCAAAATTAATGCGCAAAACCAATTAAGAGAGCGGCAATGAGTTAACAGAT 360  
 QY 361 CTAATCAATGCTTAAAGCATGACATTAGCTTCACTACAAAGCTATGCCAAGTTGAACT 420  
 DB 361 CTAATCAATGCTTAAAGCATGACATTAGCTTCACTACAAAGCTATGCCAAGTTGAACT 420  
 QY 421 AGTTTATCATCTGCTTAAAGAGCTGAAGCTGAACAGTTAACTTAATGCAACATTA 480  
 DB 421 AGTTTATCATCTGCTTAAAGAGCTGAAGCTGAACAGTTAACTTAATGCAACATTA 480  
 QY 481 GAACAACCTAAAGAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCT 540  
 DB 481 GAACAACCTAAAGAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCT 540  
 QY 541 AAAACGACTTTTGATATGAACACCAAAATTTAGTTGAAGCATCAAAAGCACTTAAAAAC 600

DB 541 AAAACGACTTTTGATATGAACATCAAAATTTAGTTGAAGCATCAAAAGCACTTAAAAAC 600  
 QY 601 ACTTTAAGCAACGCTGCTTAAAGCTTGAAGCTTGTATCAATGCTTAAATCAATTT 660  
 DB 601 ACTTTAAGCAACGCTGCTTAAAGCTTGAAGCTTGTATCAATGCTTAAATCAATTT 660  
 QY 661 CGCAATTAATTTAGTGAATCTTAAAGCTTGAAGCTTGTATCAATGCTTAAATCAATTT 720  
 DB 661 CGCAATTAATTTAGTGAATCTTAAAGCTTGAAGCTTGTATCAATGCTTAAATCAATTT 720  
 QY 721 CCCTAAATGGGGAGAGCTTTTGAATTTGAATGATTTACTCACTTAATCGAATTT 780  
 DB 721 CCCTAAATGGGGAGAGCTTTTGAATTTGAATGATTTACTCACTTAATCGAATTT 780  
 QY 781 AATAATGCTTAACTTAAATGAACAAAGCTAATGCTGATGCTTAAATGCTTAAATGCT 840  
 DB 781 AATAATGCTTAACTTAAATGAACAAAGCTAATGCTGATGCTTAAATGCTTAAATGCT 840  
 QY 841 TTTATTAATAAAGATTTCAAAATTAATGAACAAAGCTTGAAGCTTGAAGCTTGAAGCT 900  
 DB 841 TTTATTAATAAAGATTTCAAAATTAATGAACAAAGCTTGAAGCTTGAAGCTTGAAGCT 900  
 QY 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGTGATGATGATGATGATGAT 960  
 DB 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTAGTGTGATGATGATGATGATGAT 960  
 QY 961 TATTAATTAATGAACAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1020  
 DB 961 TATTAATTAATGAACAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1020  
 QY 1021 AACAGAAATGATACAGATGTTCTTGTGATTTATGATTTAGTGAAGCAACAAACAGAG 1080  
 DB 1021 AACAGAAATGATACAGATGTTCTTGTGATTTATGATTTAGTGAAGCAACAAACAGAG 1080  
 QY 1081 TACCAATTTAGTTTGAACATGATGTCATCAATGCTTATTTATTTATTTATTTATTT 1140  
 DB 1081 TACCAATTTAGTTTGAACATGATGTCATCAATGCTTATTTATTTATTTATTTATTT 1140  
 QY 1141 TTGGTTAAAGAGCTGATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAAT 1200  
 DB 1141 TTGGTTAAAGAGCTGATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAAT 1200  
 QY 1201 GTTCAACAAGTTGATGCTTCACTTCAATGCTTCAATGCTTCAATGCTTCAATGCTTCA 1260  
 DB 1201 GTTCAACAAGTTGATGCTTCACTTCAATGCTTCAATGCTTCAATGCTTCAATGCTTCA 1260  
 QY 1261 CCAGCAGTTGATGATTAAGTTGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 1306  
 DB 1261 CCAGCAGTTGATGATTAAGTTGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 1306  
 RESULT 3  
 AA053419  
 ID AA053419 standard; DNA; 1387 BP.  
 AC AA053419;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 16-JUN-1994 (first entry)  
 DE Mycoplasma gallisepticum antigen coding sequence.  
 XX  
 XX Vaccine; mycoplasma infection; poultry; fowl; ds.  
 XX Mycoplasma gallisepticum.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 202..1308  
 FT /note= "ORF includes 2 NNN codons"  
 FT misc\_difference 51



XX Mycoplasma gallisepticum antigenic protein TTM-1 gene.  
 XX Antigenic protein; vaccine; poultry; diagnosis; TTM-1, ds.  
 KW Mycoplasma gallisepticum.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 202..1308  
 FT /\*tag= a  
 XX JP07133295-A.  
 XX 23-MAY-1995.  
 XX 27-AUG-1993; 93JP-0213102.  
 XX 27-AUG-1993; 93JP-0213102.  
 XX (JAPG ) NIPPON ZEON KK.  
 XX (SHIO ) SHIONOGI & CO LTD.  
 XX MPI: 1995-220782/29.  
 XX P-PSDB; AAR76955.  
 XX A new antigenic protein which reacts with Mycoplasma gallisepticum -  
 PT is useful in a component vaccine for use against poultry infected  
 PT with M. gallisepticum.  
 XX Claim 6; Figs 5-6; 33pp; Japanese.  
 XX AA094711 encodes AAR76955 Mycoplasma gallisepticum antigenic protein  
 CC TTM-1. TTM-1 can be used as a vaccine for M. gallisepticum  
 CC infectious diseases in poultry, and as a diagnostic agent for  
 CC M. gallisepticum.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX Sequence 1387 BP; 531 A; 222 C; 201 G; 427 T; 6 other:  
 SQ  
 Query Match 96.5%; Score 1260.8; DB 16; Length 1387;  
 .Best Local Similarity 98.2%; Pred. No. 1.8e-195;  
 Matches 1282; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

DB 421 AGTTATCATCTGCTTAATAGTGAAGCTGAACAGTTAAACAATTAACCTTAATGCAACATA 480  
 QY 481 GAACAACATAAATGAGCTAAATTAATTTAGATCAGCATCAACCAAGTAATACGAT 540  
 DB 481 GAACAACATAAATGAGCTAAATTAATTTAGATCAGCATCAACCAAGTAATACGAT 540  
 QY 541 AAAACGACTTTGATTAATGAACACCAAAATTTAGTGAAGCATACAAAGCATTAACCA 600  
 DB 541 AAAACGACTTTGATTAATGAACATCAAAATTTAGTGAAGCATACCAAGCATTAACCA 600  
 QY 601 ACTTTAGAACACAGCTGCTAATCACTTGAAGGTTTGCATCACTGCTTAATCAAAAT 660  
 DB 601 ACTTTAGAACACAGCTGCTAATCACTTGAAGGTTTGCATCACTGCTTAATCAAAAT 660  
 QY 661 CGCAATTAATTTAGTGAATCTAATCAATTAAGCTAGTATTAAATCAATTAACATAGAT 720  
 DB 661 CGTAAATTAATTTAGTGAATCTAATCAATTAAGCTAGTATTAAATCAATTAACATAGAT 720  
 QY 721 CCACATAATGGGGAGCGCTTTTGAATTTGAATGAGATTACTACAGTTAATCGAATATT 780  
 DB 721 CCACATAATGGGGAGCGCTTTTGAATTTGAATGAGATTACTACAGTTAATCGAATATT 780  
 QY 781 AATATATAGCTTATCACTTATTAATGAACAAAGCTAATGCTGATTTCTAATATGT 840  
 DB 781 AATATATAGCTTATCACTTATTAATGAACAAAGCTAATGCTGATTTCTAATATAGT 840  
 QY 841 TTTATTAATAAAGATTTCAAAATTAATGAACAAAGTTTGAAGGACTTTTACAAAGCT 900  
 DB 841 TTTATTAATAAAGATTTCAAAATTAATGAACAAAGTTTGAAGGACTTTTACAAAGCT 900  
 QY 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGAATGCTGATTAACACCCGTCAT 960  
 DB 901 AATGTTCAACCTTCAAACTACAGTTTGTGCTTTTGAATGCTGATTAACACCCGTCAT 960  
 QY 961 TATTAATATGCAAGAACAGCCGTTTGAATGCTGATTAACCTTCAATGAATTTCTTGA 1020  
 DB 961 TATTAATATGCAAGAACAGCCGTTTGAATGCTGATTAACCTTCAATGAATTTCTTGA 1020  
 QY 1021 AACAGATATGATCACAGATGTTTCTGATTTATATAGTTAGCTGGAACAAACAGGAG 1080  
 DB 1021 AACAGATATGATCACAGATGTTTCTGATTTATATAGTTAGCTGGAACAAACAGGAG 1080  
 QY 1081 TACCAATTAATGTTTACCAATATATGCTCAATCAACCTGATTAATTTCCCTTAATAG 1140  
 DB 1081 TACCAATTAATGTTTACCAATATATGCTCAATCAACCTGATTAATTTCCCTTAATAG 1140  
 QY 1141 TTGCTTAAGCAGCTGATGCTAATTAACCTTGAATTAACATTAATTAATGAAT 1200  
 DB 1141 TTGCTTAAGCAGCTGATGCTAATTAACCTTGAATTAACATTAATTAATGAAT 1200  
 QY 1201 GTTCAACAGTTGATGCTGCACTTCAACTAGTGAATTAATTAATTAATTAATTAAT 1260  
 DB 1201 GTTCAACAGTTGATGCTGCACTTCAACTAGTGAATTAATTAATTAATTAATTAAT 1260  
 QY 1261 CCAGCAGTTGATGATTAATTAATGTTGCTAATTAATGTTTATCAAGGTT 1306  
 DB 1261 CCAGCAGTTGATGATTAATTAATGTTGCTAATTAATGTTTATCAAGGTT 1306  
 RESULT 5  
 AA077856  
 ID AA077856 standard; DNA; 1935 BP.  
 XX  
 XX AA077856;  
 XX 25-MAR-2003 (updated)  
 DT 23-JUN-1995 (first entry)  
 XX  
 DE Mycoplasma gallisepticum antigen TM-66 coding sequence.  
 XX recombinant avipox virus; live vaccine; mycoplasma antigen; ds.  
 KW Mycoplasma gallisepticum.  
 OS

```

XX Key Location/Qualifiers
FH CDS 54..1886
FT /tag= a
FT /product= Mycoplasma_antigen
FT mat_peptide 54..1883
FT /tag= b
FT /codon= seq:TGA,aa:Trp
XX
XX MO9423019-A1.
XX
XX 13-OCT-1994.
XX
XX 31-MAR-1994; 94MO-JP00541.
XX
XX 31-MAR-1993; 93JP-0074139.
XX 30-SEP-1993; 93JP-0245625.
XX
XX (JAPG) NIPPON ZEON KK.
XX (SHIO) SHIONOGI & CO LTD.
XX
XX Aoyama S, Punato H, Iritani Y, Ohkawa S, Ohnawa I,
XX Saeki S, Saitos, Takahashi K;
XX WPI; 1994-333181/41.
XX P-PSDB; AAR63229.
XX
XX Recombinant avipox virus combining DNA encoding a polypeptide -
XX exhibiting antigenicity of mycoplasma, useful for the production
XX of a live vaccine
XX
XX Claim 4; Page 78-81; 123pp; Japanese.
XX
XX A restriction fragment of the insert of M.gallisepticum genomic clone
XX PUM-66 containing an open reading frame was sequenced (AA077856). The
XX ORF encodes an antigenic polypeptide. A recombinant avipox virus
XX comprising the coding sequence can be used as a live vaccine to
XX protect against infection by Mycoplasma gallisepticum.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 1935 BP; 728 A; 314 C; 307 G; 586 T; 0 other;
XX
XX Query Match 83.2%; Score 1086.8; DB 15; Length 1935;
XX Best Local Similarity 96.7%; Pred. No. 2.5e-167;
XX Matches 1120; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
XX
XX 150 TTTATTTTATTTTATTTAGT CATCTTTAAGATATATATCTTAATTTCTATGAATA 208
XX 1 TTTATTTTATTTTATTTTATTTTAAATATATATATTTTAAATTTCTATGAATA 60
XX
XX 209 AGAAGAATCATCTTAAGACTATTTAGTTGTAGTAAACATCCCTTCTAGCATTTG 268
XX 61 AAAAAAGATCATCTTAAGACTATTTAGTTGTAGTAAACATCCCTTCTAGCATTTG 120
XX
XX 269 GGATTTCTAGCTGTATGTCTTATTTACTTAAAGATGCAACCAATATATGCAAAACC 328
XX 121 GGATTTCTAGCTGTATGTCTTATTTACTTAAAGATGCAACCAATATATGCAAAACC 180
XX
XX 329 AATTGAAGCAGCGCGAATGAGTTAACAAGATCTATCAATGCTAAAGCATGACATTAG 388
XX 181 AATTGAAGCAGCGCGAATGAGTTAACAAGATCTATCAATGCTAAAGCATGACATTAG 240
XX
XX 389 CTTCACTCAAGACTATGATGCAAGATTGAAGTTATTCATGCTTATATGTAAGTGTG 448
XX 241 CTTCACTCAAGACTATGATGCAAGATTGAAGTTATTCATGCTTATATGTAAGTGTG 300
XX
XX 449 AAACAGTTAACAATACCTTAATGCAACATTAGAACAACTAAATGCTTAAATCTAATT 508
XX 301 AAACAGTTAACAATACCTTAATGCAACATTAGAACAACTAAATGCTTAAATCTAATT 360
XX
XX 509 TAGAATGAGCCATCAACCAAGTATATAGGATTAAGCACTTTTGATATATGAACCCAA 568
XX 361 TAGAATGAGCCATCAACCAAGTATATAGGATTAAGCACTTTTGATATATGAACCCAA 420
XX

```

```

XX 569 ATTATGTTGAAGCATACAAAGCATTAATAAACCACTTTTGAACAACGTCTACTAACCCTTG 628
XX 421 ATTATGTTGAAGCATACAAAGCATTAATAAACCACTTTTGAACAACGTCTACTAACCCTTG 480
XX
XX 629 AAGGTTGTCATCAACGTCTTATATGAATTCGAAATATTTGTGATCTATATCAATA 688
XX 481 AAGGTTGTCATCAACGTCTTATATGAATTCGAAATATTTGTGATCTATATCAATA 540
XX
XX 689 AAGCTAGTAGTTTATATTAATAAACACTAGATCACTAAATGAGGGAACGCTTTAGATT 748
XX 541 AAGCTAGTAGTTTATATTAATAAACACTAGATCACTAAATGAGGGAACGCTTTAGATT 600
XX
XX 749 CTAATGAGATTACTACAGTTATATGGAATATTAATATAGTTATCAACTATTAATGAAAC 808
XX 601 CTAATGAGATTACTACAGTTATATGGAATATTAATATAGTTATCAACTATTAATGAAAC 660
XX
XX 809 AAAGACTAATGCTGATGCACTTATCTATATAGTTTATTAATAAAGTATCAATAATATG 868
XX 661 AAAGACTAATGCTGATGCACTTATCTATATAGTTTATTAATAAAGTATCAATAATATG 720
XX
XX 869 AAACAAAGTTTGTAGGACCTTTTCAAAACGCTAATGTTCAACCTTCAAACTACAGTTTGG 928
XX 721 AAACAAAGTTTGTAGGACCTTTTCAAAACGCTAATGTTCAACCTTCAAACTACAGTTTGG 780
XX
XX 929 TTGCTTTTATGCTGATGTAACACCCGCTCAATTTAATATGCAAGAACCGTTTGA 988
XX 781 TTGCTTTTATGCTGATGTAACACCCGCTCAATTTAATATGCAAGAACCGTTTGA 840
XX
XX 989 ATGCGATGAACTTCAAGTGAATCTTGCAACACAGAAATGATGACAGATGTTCTT 1048
XX 841 ATGCGATGAACTTCAAGTGAATCTTGCAACACAGAAATGATGATGATGTTCTT 900
XX
XX 1049 GGATTTATAGTTAGCTGGAACAAACAGAGTACCAATTTAGTTTACCAATATGATGTC 1108
XX 901 GAATTTATAGTTAGCTGGAACAAACAGAGTACCAATTTAGTTTACCAATATGATGTC 960
XX
XX 1109 CATACACTGCTTATTTATTTTCCCTTATATAGTTGTTAAAGCAGCTGATCTAATAAG 1168
XX 961 CATACACTGCTTATTTATTTTCCCTTATATAGTTGTTAAAGCAGCTGATCTAATAAG 1020
XX
XX 1169 TTGATTTACATACAAATTAATAATGAAATGTTCAACAAGTTGATGCCACTTCAA 1228
XX 1021 TTGATTTACATACAAATTAATAATGAAATGTTCAACAAGTTGATGCCACTTCAA 1080
XX
XX 1229 CTAGTGAATATATATCTACAGCTATATCCAACTCCAGAGTTGATGATTTAAAGTCTTA 1288
XX 1081 CTAGTGAATATATATCTACAGCTATATCCAACTCCAGAGTTGATGATTTAAAGTCTTA 1140
XX
XX 1289 AAATCGTTTATACAGTT 1306
XX 1141 AAATCGTTTATACAGTT 1158
XX
XX
XX RESULT 6
XX AAT96595
XX ID AAT96595 standard; DNA; 1371 BP.
XX
XX AAT96595;
XX
XX 15-JUL-1998 (first entry)
XX
XX Hybrid Marek's disease virus/M. gallisepticum gene 40 K-S.
XX
XX Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
XX antigen; vaccine; poultry; ds.
XX
XX Chimeric - Marek's disease gammaherpesvirus.
XX Chimeric - Mycoplasma gallisepticum.
XX
XX Key Location/Qualifiers
XX FH 1..1371
XX FT /tag= a
XX

```

FT	misc_feature	/product= "protein 40 K-S"
FT	1..192	
FT	/*tag= b	
FT	/note= "derived from Marek's disease virus gb gene"	
FT	193..11368	
FT	/*tag= c	
FT	/note= "derived from gene encoding M. gallisepticum antigen"	
FT		
PN	MO9736924-A1.	
PD	09-OCT-1997.	
XX		
XX	28-MAR-1997;	97WO-JP01084.
XX	29-MAR-1996;	96JP-0103548.
XX		
PA	(JAPG ) NIPPON ZEON KK.	
XX		
PI	Saito S, Tsuzaki Y, Yanagida N;	
XX		
DR	WPI, 1997-503046/46.	
DR	P-PSDB; AAM36050.	
XX		
PT	Fusion protein comprising herpes virus outer membrane protein and antigenic polypeptide - for prevention of infection by Mycoplasma gallisepticum, especially in poultry	
PS	Disclosure; Page 16-19; 51pp; Japanese.	
XX		
CC	This sequence encodes the chimeric protein 40 K-S which comprises a fragment of the Marek's disease virus outer membrane protein gb fused to an antigenic protein from Mycoplasma gallisepticum. The chimeric protein can be used in recombinant live vaccines for prevention of infection by Mycoplasma gallisepticum, especially as the outer membrane protein shows antigenicity in poultry.	
XX		
SO	Sequence 1371 BP; 491 A; 242 C; 232 G; 406 T; 0 other;	
Query Match	78.4%; Score 1024.2; DB 18; Length 1371;	
Best Local Similarity	99.7%; Pred. No. 3.4e-157;	
Matches 1026; Conservative	0; Mismatches 3; Indels 0; Gaps 0	
QY	278 GCTGTATGTCCTAATCTAATAAAGATGCACCAACCAATATAGGCCCAATTAAGG	337
DB	188 GCTGTATGTCCTAATCTAATAAAGATGCACCAACCAATATAGGCCCAATTAAGG	247
QY	338 CAGCGCGAATGGAGTTAACAAGTCTAATCAAGCTTAAGCGATGACATTAGCTTCATAC	397
DB	248 CAGCGCGAATGGAGTTAACAAGTCTAATCAAGCTTAAGCGATGACATTAGCTTCATAC	307
QY	398 AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTAAAGCTGAACAGTTA	457
DB	308 AAGACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATAGTAAAGCTGAACAGTTA	367
QY	458 ACAATTAACCTTAATGCAACATTAGAACCACTAATAATGGCTTAATACTAATTTAGATCG	517
DB	368 ACAATTAACCTTAATGCAACATTAGAACCACTAATAATGGCTTAATACTAATTTAGATCG	427
QY	518 CCATTAACAACAGCTTAATAGCGGTAATAAGCACTTTTGATTAATAACCAATTTAGTGG	577
DB	428 CCATTAACAACAGCTTAATAGCGGTAATAAGCACTTTTGATTAATAACCAATTTAGTGG	487
QY	578 AAGCATCAAAAGCACTAATAAACCCTTTAGAACCAAGTCTACTAACCCTTGAAGGTTGT	637
DB	488 AAGCATCAAAAGCACTAATAAACCCTTTAGAACCAAGTCTACTAACCCTTGAAGGTTGT	547
QY	638 CATCAACTGCTTATATACAAATTCGCAATATTTAGTGATCTATACAAATAAGCTAGTA	697
DB	548 CATCAACTGCTTATATACAAATTCGCAATATTTAGTGATCTATACAAATAAGCTAGTA	607
QY	698 GTTTAATAACCTTAACCACTAGATCCACCTTAATGGGGGAAAGCTTTTAAATTTGAATGAGA	757

Dd		608	GTTTAATACAAACAAGCTAGATCCACTAATAATGGGGAAGCGCTTGTAGATTCTAATAGA	667
Oy		758	TTACTACAGTTAATCGGAATATTAAATAATGCTTATCACTATTAAATGACAAAAGCTA	817
Dd		668	TTACTACAGCTAATAAGAAATATTAATATATACGTTATCAACTATTAAATGAACAAAAGCTA	727
Oy		818	AATGCTAGTCATATATCTAATATGTTTTTATTAATAAAGATCAAAATAATGAACAAAGTT	877
Dd		728	AATGCTAGTGCAATATATCTAATATGTTTTTATTAATAAAGATCAAAATAATGAACAAAGTT	787
Oy		878	TTTAGAGGACCTTTTACAAAACGCTATATGTTCAACCTTCAAACCTACAGCTTTTGTCTTTTA	937
Dd		788	TTTAGAGGACCTTTTACAAAACGCTATATGTTTCAACCTTCAAACCTACAGCTTTTGTCTTTTA	847
Oy		938	GTGCTGATGTAAACCCGCTCAATTAATAATAATGCAAGAAGAACCGTTGGAATGCTGATG	997
Dd		848	GTGCTGATGTAAACCCGCTCAATTAATAATAATGCAAGAAGAACCGTTGGAATGCTGATG	907
Oy		998	AACCTTCAAGTAAATTTCTTGCCAACACGAATTACTATCACAGATGTTCTTGGAATTTATA	1057
Dd		908	AACCTTCAAGTAAATTTCTTGCCAACACGAATATATCACAGATGTTCTTGGAATTTATA	967
Oy		1058	GTTTAGCTGGAACAAACAGAAAGTACCATTATTAGTTTAAAGCACTATGTCATCAACTG	1117
Dd		968	GTTTAGCTGGAACAAACAGAAAGTACCATTATTAGTTTAAAGCACTATGTCATCAACTG	1027
Oy		1118	GTTAATTATATATTCCTTATTAAGTTGTTTAAAGCAGTGATGCTAATPAAGCTTGGAATTAC	1177
Dd		1028	GTTAATTATATATTCCTTATTAAGTTGTTTAAAGCAGTGATGCTAATPAAGCTTGGAATTAC	1087
Oy		1178	AATACAAATTAAATTAATGGAATATGTTCAACAGTTGAGTTGCCACTTCAACTAGTGCA	1237
Dd		1088	AATACAAATTAAATTAATGGAATATGTTCAACAGTTGAGTTGCCACTTCAACTAGTGCA	1147
Oy		1238	ATAATACTACAGCTAATCCAACTCCAGACGTTGATGAGATTAAAGTTGCTAAAAATCGTTT	1297
Dd		1148	ATAATACTACAGCTAATCCAACTCCAGACGTTGATGAGATTAAAGTTGCTAAAAATCGTTT	1207
Oy		1298	TATCAGGTT 1306	
Dd		1208	TATCAGGTT 1216	
 RESULT 7 AAT96596 ID AAT96596 standard; DNA; 3261 BP. XX AC AAT96596; XX DT 15-JUN-1998 (first entry) XX DE Hybrid Marek's disease virus/M. gallisepticum gene 40 K-C. XX KM Chimeric; Marek's disease virus; outer membrane protein; fusion protein; XX antigen; vaccine; poultry; ds. OS Chimeric - Marek's disease gammaherpesvirus. OS Chimeric - Mycoplasma gallisepticum. XX PH Key Location/Qualifiers FT CDS 1..3261 FT /*tag= a /product= "protein 40 K-C" FT 1..2016 FT misc_feature /tag= b /note= "derived from Marek's disease virus gb gene" FT 2077..3258 FT /*tag= c FT /note= "derived from gene encoding M. gallisepticum antigen" XX PN W09736924-A1 XX				



PD 09-OCT-1997.  
 XX 28-MAR-1997; 97MO-JP01084.  
 XX 29-MAR-1996; 96JP-0103548.  
 PR (JAPG ) NIPPON ZEON KK.  
 PA Saito S, Tsuzaki Y, Yanagida N;  
 PI MPI; 1997-503046/46.  
 DR P-PSDB; AAM36051.  
 XX  
 XX Fusion protein comprising herpes virus outer membrane protein and  
 PT antigenic polypeptide for prevention of infection by Mycoplasma  
 PT gallisepticum, especially in poultry  
 XX  
 XX Disclosure; Page 22-30; 51pp; Japanese.  
 XX  
 CC This sequence encodes the chimeric protein 40 K-C which comprises a  
 CC fragment of the Marek's disease virus outer membrane protein 9B fused  
 CC to an antigenic protein from Mycoplasma gallisepticum. The chimeric  
 CC protein can be used in recombinant live vaccines for prevention of  
 CC infection by Mycoplasma gallisepticum, especially as the outer membrane  
 CC protein shows antigenicity in poultry.  
 XX  
 XX Sequence 3261 BP; 1099 A; 617 C; 648 G; 897 T; 0 other;  
 SQ  
 \* Query Match 78.3%; Score 1022.6; DB 18; Length 3261;  
 Best Local Similarity 99.6%; Pred. No. 5.9e-157;  
 Matches 1025; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 278 GCTGATGCTCTTATTAATAAAGATGCAACCCCAATATAGCCCAACCAATTAGAG 337  
 Db 2078 GCTGATGCTCTTATTAATAAAGATGCAACCCCAATATAGCCCAACCAATTAGAG 2137  
 QY 338 CAGCCGGAATGAGTAAACAGATCTATCATGCTAAAGCGATGACATTAGCTTACATAC 397  
 Db 2138 CAGCCGGAATGAGTAAACAGATCTATCATGCTAAAGCGATGACATTAGCTTACATAC 2197  
 QY 398 AAGACTATGCCAAGATTGAAGTATTATCATCTGCTTATAGTGAAGCTGAACAGTTA 457  
 Db 2198 AAGACTATGCCAAGATTGAAGTATTATCATCTGCTTATAGTGAAGCTGAACAGTTA 2257  
 QY 458 ACAATATACCTTATGCAATAGAAACAATAAATGCTAAATCTAATTTAGATCG 517  
 Db 2258 ACAATATACCTTATGCAATAGAAACAATAAATGCTAAATCTAATTTAGATCG 2317  
 QY 518 CCATCAACCAAGCTATATGAGTAAACAGACTTTGATATGAAACCCCAATTAGTTG 577  
 Db 2318 CCATCAACCAAGCTATATGAGTAAACAGACTTTGATATGAAACCCCAATTAGTTG 2377  
 QY 578 AAGCATACAAAGCATTAATAAACCATTGGAACAACGTCCTAATCTTGAAGTTGT 637  
 Db 2378 AAGCATACAAAGCATTAATAAACCATTGGAACAACGTCCTAATCTTGAAGTTGT 2437  
 QY 638 CATCACTGCTATATATCAATTCGCAATATTTGTGGATGTATACATAAAGCTAGTA 697  
 Db 2438 CATCACTGCTATATATCAATTCGCAATATTTGTGGATGTATACATAAAGCTAGTA 2497  
 QY 698 GTTATATATCTAAACACTAGATCCACTAAATGCGGGAACGCTTTTATAGTTCTATGAGA 757  
 Db 2498 GTTATATATCTAAACACTAGATCCACTAAATGCGGGAACGCTTTTATAGTTCTATGAGA 2557  
 QY 758 TTACTACAGTTATCGGAATTTAATATAGTTTCAACTATTAATGACAAAGACTA 817  
 Db 2558 TTACTACAGTTATCGGAATTTAATATAGTTTCAACTATTAATGACAAAGACTA 2617  
 QY 818 ATGCTGAGCATTTATCTATATAGTTTATTAATAAAGTATCAAAATATAGAACAAAGT 877  
 Db 2618 ATGCTGAGCATTTATCTATATAGTTTATTAATAAAGTATCAAAATATAGAACAAAGT 2677  
 QY 878 TTGTAGGAGCTTTTACAAACGCTAAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTA 937

Db 2678 TTGTAGGAGCTTTTACAAACGCTAAATGTTCAACCTTCAAACTACAGTTTGTGCTTTTA 2737  
 QY 938 GTGCTGATGTAAACCCGCTCAATTATTAATATGCAAGAAGACCGTTTGAATGTGATG 997  
 Db 2738 GTGCTGATGTAAACCCGCTCAATTATTAATATGCAAGAAGACCGTTTGAATGTGATG 2797  
 QY 998 AACCTTCAAGTAAATTTCTTGCAACACGAATAGTATCAAGATGTTTCTTGATTTATA 1057  
 Db 2798 AACCTTCAAGTAAATTTCTTGCAACACGAATAGTATCAAGATGTTTCTTGATTTATA 2857  
 QY 1058 GTTATGCTGAACAACAACGAGTACCAATTAGTTTATAGCAATGATGTCATCAACTG 1117  
 Db 2858 GTTATGCTGAACAACAACGAGTACCAATTAGTTTATAGCAATGATGTCATCAACTG 2917  
 QY 1118 GTTATTTATTTCCCTTATTAATGTTGTTAAAGCAGTATGCTAATACGTTGATTAAC 1177  
 Db 2918 GTTATTTATTTCCCTTATTAATGTTGTTAAAGCAGTATGCTAATACGTTGATTAAC 2977  
 QY 1178 AATACAAATTAATATATGGAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTCAAA 1237  
 Db 2978 AATACAAATTAATATATGGAATGTTCAACAAGTTGAGTTGCCACTTCAACTAGTCAAA 3037  
 QY 1238 ATAATACTACAGCTATATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTT 1297  
 Db 3038 ATAATACTACAGCTATATCCAACTCCAGCAGTTGATGAGATTAAAGTTGCTAAATCGTTT 3097  
 QY 1298 TATCAGGTT 1306  
 Db 3098 TATCAGGTT 3106

RESULT 8  
 AAT04076  
 ID AAT04076 standard; DNA; 1015 BP.  
 XX  
 XX AAT04076;  
 AC  
 XX  
 DT 19-JUL-1996 (first entry)  
 XX  
 DE M.gallisepticum DNA sequence II encodes 261 amino acid protein.  
 XX  
 KW Detection; probe; primer; PCR; amplification; secretion; lung;  
 XX avian chronic respiratory disease; respiratory tract; nasal cavity; ds.  
 OS Mycoplasma gallisepticum.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 202..987  
 FT /\*tag= a  
 XX  
 PN JP07236498-A.  
 XX  
 PD 12-SEP-1995.  
 XX  
 PF 25-FEB-1994; 94JP-0052764.  
 XX  
 PR 25-FEB-1994; 94JP-0052764.  
 XX  
 PA (JAPG ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX  
 DR MPI; 1995-347462/45.  
 DR P-PSDB; AAR79911.  
 XX  
 PT Detection of Mycoplasma gallisepticum - for the quick detection,  
 PT i.e. within one day, of avian chronic respiratory diseases  
 XX  
 PS Claim 3; Page 10-11; 11pp; Japanese.  
 XX  
 CC This is the nucleotide sequence of a fragment of the Mycoplasma  
 CC gallisepticum genome which codes for a 261 amino acid protein.  
 CC This sequence and the sequence of AAT04075 (encoding a 661 amino



CC acid protein) can be used to detect *M.gallisepticum* using probes  
 CC based on nucleotides 1125-1648 and primers based on nucleotides  
 CC 449-466, the complement of bases 893-919, 1908-1934 and the  
 CC complement of bases 2184-2210 of AAT04075, and a probe based on  
 CC nucleotide 718-41 of this sequence. The method using these sequences  
 CC is faster i.e. is able to detect *M.gallisepticum*, which causes avian  
 CC chronic respiratory diseases, within one day, from avian secretions,  
 CC washings from the lung, respiratory tract, nasal cavity, etc.  
 CC  
 XX

Sequence 1015 BP; 400 A; 162 C; 138 G; 315 T; 0 other;

Query Match 75.8%; Score 989.4; DB 16; Length 1015;  
 Best Local Similarity 98.4%; Pred. No. 1.4e-151;  
 Matches 999; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

Oy 1 AAAAAACATGATGTTGTAATCTGATATCTTGTAAAAAACAATAATCTTCAACA 60
Db 1 AAAAAACATGATGTTGTAATCTGATATCTTGTAAAAAACAATAATCTTCAACA 60
Oy 61 AATCTTAATTAATAGCCGTTAAATTAATAAAAAATGATTTTCTTATC 120
Db 61 AATCTTAATTAATAGCCGTTAAATTAATAAAAAATGATTTTCTTATC 120
Oy 121 AACCAAAATCTCTAGTAATAAAGCCTATTATTTTATTTTATGATCTTTAAGAT 180
Db 121 AACCAAAATCTCTAGTAATAAAGCCTATTATTTTATTTTATGATCTTTAAGAT 180
Oy 181 ATAAATATATCTTAATATTTCTATGATGAAGAAAGAAATCAATCTTAAAGATTTG 240
Db 181 ATAAATATATCTTAATATTTCTATGATGAAGAAAGAAATCAATCTTAAAGATTTG 240
Oy 241 TTAGGTACACATCTTCTTCTAGCATTTGGGATTTCTAGCTGATGTCTATTACTAA 300
Db 241 TTAGGTACACATCTTCTTCTAGCATTTGGGATTTCTAGCTGATGTCTATTACTAA 300
Oy 241 TTAGGTACACATCTTCTTCTAGCATTTGGGATTTCTAGCTGATGTCTATTACTAA 300
Db 241 TTAGGTACACATCTTCTTCTAGCATTTGGGATTTCTAGCTGATGTCTATTACTAA 300
Oy 301 GATGCAAAACCAATATATGCGCAAAACCAATTTAGAAGACGCGAATGAGTTAAGAT 360
Db 301 GATGCAAAACCAATATATGCGCAAAACCAATTTAGAAGACGCGAATGAGTTAAGAT 360
Oy 301 GATGCAAAACCAATATATGCGCAAAACCAATTTAGAAGACGCGAATGAGTTAAGAT 360
Db 301 GATGCAAAACCAATATATGCGCAAAACCAATTTAGAAGACGCGAATGAGTTAAGAT 360
Oy 361 CTAATCAATGCTTAAGCCATGATTAATCTTCACTCAACAAGATATGCCAATTTAAGT 420
Db 361 CTAATCAATGCTTAAGCCATGATTAATCTTCACTCAACAAGATATGCCAATTTAAGT 420
Oy 361 CTAATCAATGCTTAAGCCATGATTAATCTTCACTCAACAAGATATGCCAATTTAAGT 420
Db 361 CTAATCAATGCTTAAGCCATGATTAATCTTCACTCAACAAGATATGCCAATTTAAGT 420
Oy 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACTTAATGCAACATTA 480
Db 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACTTAATGCAACATTA 480
Oy 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACTTAATGCAACATTA 480
Db 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACTTAATGCAACATTA 480
Oy 481 GAACAACATAAATGGCTAATAATTAATTAAGATCAGCATCAACCAAGCTAATCGAT 540
Db 481 GAACAACATAAATGGCTAATAATTAATTAAGATCAGCATCAACCAAGCTAATCGAT 540
Oy 481 GAACAACATAAATGGCTAATAATTAATTAAGATCAGCATCAACCAAGCTAATCGAT 540
Db 481 GAACAACATAAATGGCTAATAATTAATTAAGATCAGCATCAACCAAGCTAATCGAT 540
Oy 541 AAAAGACTTTTGAATGAACCAATCAATTAATTAAGATCAACCAAGCTAATCGAT 600
Db 541 AAAAGACTTTTGAATGAACCAATCAATTAATTAAGATCAACCAAGCTAATCGAT 600
Oy 541 AAAAGACTTTTGAATGAACCAATCAATTAATTAAGATCAACCAAGCTAATCGAT 600
Db 541 AAAAGACTTTTGAATGAACCAATCAATTAATTAAGATCAACCAAGCTAATCGAT 600
Oy 601 ACTTAGAACAACGCTACTAATCAATTAATTAAGATCAATCAATCAATCAATCAAT 660
Db 601 ACTTAGAACAACGCTACTAATCAATTAATTAAGATCAATCAATCAATCAATCAAT 660
Oy 601 ACTTAGAACAACGCTACTAATCAATTAATTAAGATCAATCAATCAATCAATCAAT 660
Db 601 ACTTAGAACAACGCTACTAATCAATTAATTAAGATCAATCAATCAATCAATCAAT 660
Oy 661 CGCAATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAAT 720
Db 661 CGCAATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAAT 720
Oy 661 CGCAATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAAT 720
Db 661 CGCAATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAAT 720
Oy 721 CCATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAATCAAT 780
Db 721 CCATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAATCAAT 780
Oy 721 CCATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAATCAAT 780
Db 721 CCATTAATTAAGTGAATCTAATCAATTAATTAAGTGAATCTAATCAATCAATCAAT 780
Oy 781 AATAATAGCTTATCACTATTAATGAACAAAGATTAATGATGATGATGATGATGAT 840
Db 781 AATAATAGCTTATCACTATTAATGAACAAAGATTAATGATGATGATGATGATGATGAT 840
Oy 781 AATAATAGCTTATCACTATTAATGAACAAAGATTAATGATGATGATGATGATGATGAT 840
Db 781 AATAATAGCTTATCACTATTAATGAACAAAGATTAATGATGATGATGATGATGATGAT 840
Oy 841 TTTATTAATAAAGTATCAAAATATGAACAAAGTTTGTAGGAGCTTTTCAACACGCT 900
Db 841 TTTATTAATAAAGTATCAAAATATGAACAAAGTTTGTAGGAGCTTTTCAACACGCT 900

```

```

Db 841 TTTATTAATAAAGTATCAAAATATGAACAAAGTTTGTAGGAGCTTTTCAACACGCT 900
Oy 901 AATGTTCAACCTTAATCACTAATGATTTGCTTTAGTGTGATGATGATGATGATGAT 960
Db 901 AATGTTCAACCTTAATCACTAATGATTTGCTTTAGTGTGATGATGATGATGATGAT 960
Oy 961 TATTAATATGCAAGAGACCGTTTGAATGATGATGATGATGATGATGATGATGATGAT 1015
Db 961 TATTAATATGCAAGAGACCGTTTGAATGATGATGATGATGATGATGATGATGATGAT 1015

```

# RESULT 9

```

AA004687
ID AA004687 standard; DNA; 853 BP.
XX
AC AA004687;
XX
DT 10-MAR-2003 (updated)
DT 08-OCT-1990 (first entry)
XX
DE DNA encoding TMG-1 antigen.
XX
KW Mycoplasma gallisepticum; poultry; vaccine; ss.
XX
OS Mycoplasma gallisepticum.
XX
PN JF02111795-A.
XX
PD 24-APR-1990.
XX
PF 02-JUN-1989; 89JP-0136343.
XX
PR 02-JUN-1989; 89JP-0136343.
XX
PA (JAPG ) NIPPON ZEON KK.
PA (SHIO ) SHIONOGI KK.
XX
DR WPI; 1990-169109/22.
XX
P-PSDB; AAR05082.
XX
XX
XX Diagnostic and vaccine for poultry mycoplasma serum - utilises
XX antigen protein of the disease and recombinant vector
XX incorporated with its coding gene.
XX
XX Claim 5; Fig 2; 20pp; Japanese.
XX
XX The DNA can be inserted into an expression vector for the prodn. of
XX polypeptide which elicits an antigen-antibody reaction with anti-
XX mycoplasma gallisepticum poultry sera. It may also be ligated to
XX other DNA to produce fusion proteins with an N-terminal bacterial
XX enzyme sequence.
XX See also AA004686 and AA005649-53.
XX (Updated on 10-MAR-2003 to add missing OS field.)
XX
SQ Sequence 853 BP; 329 A; 138 C; 128 G; 258 T; 0 other;

Query Match 63.1%; Score 824.2; DB 11; Length 853;
Best Local Similarity 97.9%; Pred. No. 7.7e-125;
Matches 835; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Oy 163 TTAGTCATCTTTTAAAGATATATATATCTTAATATCTTAATGATTAAGAAAGATCATC 222
Db 1 TTAGTCATCTTTTAAAGATATATATATCTTAATATCTTAATGATTAAGAAAGATCATC 222
Oy 223 TTAAGAATAATTAAGTGTGATGATCAACATCTTCTTGAAGATTTGATGATGAT 282
Db 223 TTAAGAATAATTAAGTGTGATGATCAACATCTTCTTGAAGATTTGATGATGAT 282
Oy 61 TTAAGAATAATTAAGTGTGATGATCAACATCTTCTTGAAGATTTGATGATGAT 120
Db 61 TTAAGAATAATTAAGTGTGATGATCAACATCTTCTTGAAGATTTGATGATGAT 120
Oy 283 ATGCTATTAATTAAGATGATCAACATCTTCTTGAAGATTTGATGATGATGATGAT 342
Db 283 ATGCTATTAATTAAGATGATCAACATCTTCTTGAAGATTTGATGATGATGATGAT 342
Oy 121 ATGCTATTAATTAAGATGATCAACATCTTCTTGAAGATTTGATGATGATGATGAT 180
Db 121 ATGCTATTAATTAAGATGATCAACATCTTCTTGAAGATTTGATGATGATGATGAT 180
Oy 343 CGAATGAGTTAAGATGATCAATCAATGATTAAGATGATGATGATGATGATGATGAT 402

```

```

Db      181 CGAATGAGTTAATGATCTATATCAATGCTAAAGCAAGGACATTAAGCTTACCTACCAAGAC 240
Qy      403 TATGCCAAGATTGAAGCTAGTTATATCACTGCTTATATAGTAAGCTGAACAGTTAACAAT 462
Db      241 TATGCTAAAGATTGAAGCTAGTTATATCACTGCTTATATAGTAAGCTGAACAGTTAACAAT 300
Qy      463 AACCTTAATGCACTTAGAACAACATAAAATGGCTAAATCTAATTTAGATAGGCATC 522
Db      301 AACCTTAATGCACTTAGAACAACATAAAATGGCTAAATCTAATTTAGATAGGCATC 360
Qy      523 AACCAAGCTAATACGATATAAAACGATTTGATATGAAACACCAATTTAGTTGAAGA 582
Db      361 AACCAAGCTAATACGATATAAAACGATTTGATATGAAACATCAATTTAGTTGAAGA 420
Qy      583 TACAAAGCACTAAAACCACTTTAGAACAAAGCTGCTACTAACCCTGAAGGTTGTCATCA 642
Db      421 TACAAAGCACTAAAACCACTTTAGAACAAAGCTGCTACTAACCCTGAAGGTTGTCATCA 480
Qy      643 ACTGCTTATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 702
Db      481 ACTGCTTATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 540
Qy      703 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 762
Db      541 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 600
Qy      763 ACGATTAATCGGAATTTAATAATAGTTATCACTAATTAAGAACAAAAGCTAATGCT 822
Db      601 ACGATTAATCGGAATTTAATAATAGTTATCACTAATTAAGAACAAAAGCTAATGCT 660
Qy      823 GATGCTATTAATTAATAGTTTATTAATAAAGTGAATCAAAATTAAGAACAAAGTTTGTGA 882
Db      661 GATGCTATTAATTAATAGTTTATTAATAAAGTGAATCAAAATTAAGAACAAAGTTTGTGA 720
Qy      883 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTAAGTTTGTGCTTTAGTGT 942
Db      721 GGGACTTTTACAAACGCTAATGTTCAACCTTCAAACTAAGTTTGTGCTTTAGTGT 780
Qy      943 GATGTAACACCGGTCAATTAATAATTAAGAACAGACCGTTTGAATGATGAACCT 1002
Db      781 GATGTAACACCGGTCAATTAATAATTAAGAACAGACCGTTTGAATGATGAACCT 840
Qy      1003 TCAAGTAGAATTC 1015
Db      841 TCAAGTAGAATTC 853

RESULT 10
ID      AAN92574 standard; DNA; 853 BP.
XX      AAN92574;
AC      AAN92574;
XX      25-MAR-2003 (updated)
DT      11-MAY-1990 (first entry)
XX      DNA sequence of TM-1 encoding polypeptide TMG-1.
DE      Mycoplasma gallisepticum; Poultry vaccine; ss;
XX      Mycoplasma gallisepticum.
OS      Mycoplasma gallisepticum.
XX      Key      Location/Qualifiers
FH      CDS      40...822
FT      /tag a
XX      EP345021-A.
XX      06-DEC-1989.
XX      31-MAY-1989; 89EP-0305441.
XX      31-MAY-1989; 89EP-0305441.
PF

```

```

XX      02-JUN-1989; 89JP-0136343.
PR      XX
XX      (JAPG) NIPPON ZEON KK.
XX      (SHIO) SHIONOGI SEIYAKU KK.
XX      (JAPG) NIPPON ZEON KK.
XX      (SHIO) SHIONOGI SEIYAKU KK.
PI      Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
PI      Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;
DR      WPI: 1989-358393/49.
XX      P-PSDB: AAP93959.
XX      New antigenic proteins of Mycoplasma gallisepticum - useful as poultry
PT      vaccines.
PS      Disclosure; Fig.2; 31pp; English.
XX      This base sequence of TM-1 encodes the TMG-1 polypeptide which has the
CC      same sequence as that of a polypeptide expressed in Mycoplasma
CC      gallisepticum in nature. When inserted into a recombinant vector used to
CC      transform a host the antigen protein produced can be used as a vaccine to
CC      prevent and diagnose MG infection.
CC      (Updated on 25-MAR-2003 to correct PF field.)
CC      (Updated on 25-MAR-2003 to correct PA field.)
CC      (Updated on 25-MAR-2003 to correct PI field.)
SQ      Sequence 853 BP; 330 A; 139 C; 128 G; 256 T; 0 other;

Query Match      61.4%; Score 801.8; DB 10; Length 853;
Best Local Similarity 96.2%; Pred. No. 3.2e-121;
Matches 821; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy      163 TTAGTCAATCTTTAAGATATATATATATCTTAATATCTAATGATATGAATAAGAAAGATCATC 222
Db      1 TTAGTCAATCTTTAAGATATATATATATCTTAATATCTAATGATATGAATAAGAAAGATCATC 60
Qy      223 TTAAGAATTAATAGTTTGTGTTAGTACAAATCCTTTCTTACATTTGGATTTCTAGCTGT 282
Db      61 TTAAGAATTAATAGTTTGTGTTAGTACAAATCCTTTCTTACATTTGGATTTCTAGCTGT 120
Qy      283 ATGCTATTTCTAATAAAGATGCAAAACCAATTAATGAGCAAAACCAATTAAGAACGACG 342
Db      121 ATGCTATTTCTAATAAAGATGCAAAACCAATTAATGAGCAAAACCAATTAAGAACGACG 180
Qy      343 CGAATGAGTTAACAGATCTAATCAATGCTAAGCGATGAATAGCTTCACTACACAGAC 402
Db      181 CGAATGAGTTAACAGATCTAATCAATGCTAAGCGATGAATAGCTTCACTACACAGAC 240
Qy      403 TATGCCAAGATTGAAGCTAGTTATATCACTGCTTATATAGTAAGCTGAACAGTTAACAAT 462
Db      241 TATGCCAAGATTGAAGCTAGTTATATCACTGCTTATATAGTAAGCTGAACAGTTAACAAT 300
Qy      463 AACCTTAATGCACTTAGAACAACATAAAATGGCTAAATCTAATTTAGATAGGCATC 522
Db      301 AACCTTAATGCACTTAGAACAACATAAAATGGCTAAATCTAATTTAGATAGGCATC 360
Qy      523 AACCAAGCTAATACGATATAAAACGATTTGATATGAAACACCAATTTAGTTGAAGA 582
Db      361 AACCAAGCTAATACGATATAAAACGATTTGATATGAAACATCAATTTAGTTGAAGA 420
Qy      583 TACAAAGCACTAAAACCACTTTAGAACAAAGCTGCTACTAACCCTGAAGGTTGTCATCA 642
Db      421 TACAAAGCACTAAAACCACTTTAGAACAAAGCTGCTACTAACCCTGAAGGTTGTCATCA 480
Qy      643 ACTGCTTATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 702
Db      481 ACTGCTTATATCAAAATTCGCAATATTTAGTGAATCTATACATTAAGCTAGTTA 540
Qy      703 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 762
Db      541 ATTAATAAAACACTAGATCCACTAAATGGGGGAACGCTTTAGATTCTAATGAGATTACT 600

```

Qy 763 ACAGTTAATCGAATATTATTAATACGTTATCACTATTAATGAACAAAGACTAATGCT 822  
Db 601 ACAGTTAATCGAATATTATTAATACGTTATCACTATTAATGAACAAAGACTAATGCT 660  
Qy 823 GATGATATTATTAATAGTTTATTAATGAAGATGATCAAAATTAATGAACAAAGTTTGT 882  
Db 661 GATGATATTATTAATAGTTTATTAATGAAGATGATCAAAATTAATGAACAAAGTTTGT 720  
Qy 883 GGGACTTTTACAAACGCTAATGTTCAACCTCAAACTACGTTTGTGCTTTAGTCT 942  
Db 721 GGGACTTTTACAAACGCTAATGTTCAACCTCAAACTACGTTTGTGCTTTAGTCT 780  
Qy 943 GATGTAACCCGCTCAATTAATTAATGCAAGAGACCGTTGGATGATGAACCT 1002  
Db 781 GATGTAACCCGCTCAATTAATTAATGCAAGAGACCGTTGGATGATGAACCT 840  
Qy 1003 TCAAGTAGAATTC 1015  
Db 841 TCAAGTAGAATTC 853

## RESULT 11

ID AAN92568 standard; DNA; 708 BP.

XX AAN92568;

XX 25-MAR-2003 (updated)

DT 11-MAY-1990 (first entry)

DE DNA sequence of M1 encoding Mycoplasma gallisepticum (MG-1) polypeptide.

KM Mycoplasma gallisepticum; Poultry vaccine; ss;

XX Mycoplasma gallisepticum.

XX EP345021-A.

XX 06-DEC-1989.

XX 31-MAY-1989; 89EP-0305441.

PF 31-MAY-1989; 89EP-0305441.

XX 02-JUN-1988; 88JP-0136343.

XX (JAPG) NIPPON ZEON KK.

PA (SHIO) SHIONOGI SEIYAKU KK.

PA (JAPG) NIPPON ZEON KK.

PA (SHIO) SHIONOGI SEIYAKU KK.

PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

PI Kodama K, Saito S, Yanagida N, Kamogawa K, Iritani Y, Aoyama S;

DR MPI; 1989-358393/49.

DR P-PSDB; AAP93646.

XX New antigenic proteins of Mycoplasma gallisepticum - useful as poultry

XX vaccines.

XX Disclousure; Fig.1a; 31pp; English.

XX This base sequence of M1 encodes the MG1 polypeptide which elicits an

XX antigen-antibody reaction with anti-MG poultry sera. When inserted

XX into a recombinant vector used to transform a host the antigen protein

XX produced can be used as a vaccine to prevent and diagnose MG infection.

XX (Updated on 25-MAR-2003 to correct PF field.)

XX (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 708 BP; 278 A; 119 C; 107 G; 204 T; 0 other;

Query Match 52.1%; Score 679.8; DB 10; Length 708;

Best Local Similarity 97.6%; Pred. No. 1,86-101; Matches 690; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 280 TGTATGCTATTACTAAAAAGATGCAACCCAAATATGCGCAACCCAAATTTGAAGCA 339  
Db 1 TGTATGCTATTACTAAAAAGATGCAACCCAAATATGCGCAACCCAAATTTGAAGCA 60  
Qy 340 GCGGAATGAGATTAAAGATCTAATCAATGCTAAAGCATGATGATGCTTCACTA 399  
Db 61 GCGGAATGAGATTAAAGATCTAATCAATGCTAAAGCATGATGATGCTTCACTA 120  
Qy 400 GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATGATGAGCTGAAACGTTAAC 459  
Db 121 GACTATGCCAAGATTGAAGCTAGTTTATCATCTGCTTATGATGAGCTGAAACGTTAAC 180  
Qy 460 AATTAACCTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 519  
Db 181 AATTAACCTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 240  
Qy 520 ATCAACCAAGCTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 579  
Db 241 ATCAACCAAGCTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 300  
Qy 580 GCATACAAAGCACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 639  
Db 301 GCATACAAAGCACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 360  
Qy 640 TCAACTGCTTATTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 699  
Db 361 TCAACTGCTTATTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACT 420  
Qy 700 TTAATACTAAACACTAGATCCACTAATGCGGGAACGCTTTAGATTCTAATGAGATT 759  
Db 421 TTAATACTAAACACTAGATCCACTAATGCGGGAACGCTTTAGATTCTAATGAGATT 480  
Qy 760 ACTACAGTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAAT 819  
Db 481 ACTACAGTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAATGCAACTTAAT 540  
Qy 820 GCTGATGATTAATTAATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 879  
Db 541 GCTGATGATTAATTAATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600  
Qy 880 GTAGGACTTTTACAAACGCTAATGTTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 939  
Db 601 GTAGGACTTTTACAAACGCTAATGTTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCT 660  
Qy 940 GCTGATGTAACACCCGCTCAATTAATTAATGCAAGAGACCGTTTG 986  
Db 661 GCTGATGTAACACCCGCTCAATTAATTAATGCAAGAGACCGTTTG 707

## RESULT 12

ID AAO04686 standard; DNA; 708 BP.

XX AAO04686;

XX 10-MAR-2003 (updated)

DT 08-OCT-1990 (first entry)

DE DNA encoding MG-1 antigen.

KM Mycoplasma gallisepticum; poultry; vaccine; ss.

XX Mycoplasma gallisepticum.

XX JP02111795-A.

XX 24-APR-1990.

XX 02-JUN-1989; 89JP-0136343.

PR 02-JUN-1989; 89JP-0136343.  
 XX (JAPC ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI KK.  
 DR MPI; 1990-169109/22.  
 DR P-PSDB; AAR05081.  
 XX  
 PT Diagnostic and vaccine for poultry mycoplasma serum - utilizes  
 PT antigen protein of the disease and recombinant vector  
 PT incorporated with its coding gene.  
 PS  
 PS Claim 3; Fig 1a; 20pp; Japanese.  
 CC The DNA can be inserted into an expression vector for the prodn. of  
 CC polypeptide which elicits an antigen-antibody reaction with anti-  
 CC mycoplasma gallisepticum poultry sera. It may also be ligated to  
 CC other DNA to produce fusion proteins with an N-terminal bacterial  
 CC enzyme sequence.  
 CC See also AA004687 and AA005649-53.  
 CC (Updated on 10-MAR-2003 to add missing OS field.)  
 CC  
 XX Sequence 708 BP; 279 A; 119 C; 107 G; 203 T; 0 other;  
 SQ  
 Query Match 51.9%; Score 678.2; DB 11; Length 708;  
 Best Local Similarity 97.5%; Pred. No. 3.2e-101;  
 Matches 689; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 280 TGTATGCTATTACTTAAAGAGTGAACCCAAATATATGAGCCAAATAGAGCA 339  
 DB 1 TGTATGCTATTACTTAAAGAGTGAACCCAAATATATGAGCCAAATAGAGCA 60  
 QY 340 GGGCAATGAGTAAACAGATCTATCATGCTAAAGCATACATAGCTTACTTACAA 399  
 DB 61 GGGCAATGAGTAAACAGATCTATCATGCTAAAGCATACATAGCTTACTTACAA 120  
 QY 400 GACTATGCCAAGTGAAGCTAGTTATCATCTGCTTATAGAGCTGAACAGTTAAC 459  
 DB 121 GACTATGCCAAGTGAAGCTAGTTATCATCTGCTTATAGAGCTGAACAGTTAAC 180  
 QY 460 AATAACCTTATGCAACATTAGAACACTAAATATGCTAAATTTAGAAATCAGCC 519  
 DB 181 AATAACCTTATGCAACACTAGAACACTAAATATGCTAAATTTAGAAATCAGCC 240  
 QY 520 ATCAACCAAGCTAATACGATTAACGACTTTGTAATAGAACACCAATTTAGTAA 579  
 DB 241 ATCAACCAAGCTAATACGATTAACGACTTTGTAATAGAACACCAATTTAGTAA 300  
 QY 580 GCATCAAAAGCACTAAAAACAATTAGAACAGCTGCTACTACCTTGAAGTTGTCA 639  
 DB 301 GCATCAAAAGCACTAAAAACAATTAGAACAGCTGCTACTACCTTGAAGTTGTCA 360  
 QY 640 TCAACTGCTTATATCAAAATTTGCAATATTTAGTGAATCTATACATTAAGT 699  
 DB 361 TCAACTGCTTATATCAAAATTTGCAATATTTAGTGAATCTATACATTAAGT 420  
 QY 700 TTAATTAACCTAAACCTAGATCCACTAAATGGGGAAAGCTTTGATTTCTAATAG 759  
 DB 421 TTAATTAACCTAAACCTAGATCCACTAAATGGGGAAAGCTTTGATTTCTAATAG 480  
 QY 760 ACTACAGTTAATCGAATATTTAATATAGTTATCACTATTAATGAACAAAGCTAAT 819  
 DB 481 ACTACAGTTAATCGAATATTTAATATAGTTATCACTATTAATGAACAAAGCTAAT 540  
 QY 820 GCTATGCTATTTCTAATATGTTTATTAATAAAGTGAATCAAAATATGAACAAAGTTT 879  
 DB 541 GCTATGCTATTTCTAATATGTTTATTAATAAAGTGAATCAAAATATGAACAAAGTTT 600  
 QY 880 GTAGGACCTTTTACAAAGCTAATGCTCAACCTTCAAGTTGGTGGCTTTAGT 939  
 DB 601 GTAGGACCTTTTACAAAGCTAATGCTCAACCTTCAAGTTGGTGGCTTTAGT 660  
 QY 940 GCTGATGTACACCCGTCATTAATTAATATGCAAGAGACCGTTTG 986

DB 661 GCTGATGTACACCCGTCATTAATTAATATGCAAGAGAACCGTTTG 707  
 |||  
 RESULT 13  
 ID AAT04075 standard; DNA; 2196 BP.  
 AC AAT04075;  
 AC 19-JUL-1996 (first entry)  
 DE M.gallisepticum DNA sequence I encodes 661 amino acid protein.  
 XX  
 XX Detection; probe; primer; PCR; amplification; secretion; lung;  
 KW avian chronic respiratory disease; respiratory tract; nasal cavity; ds.  
 OS Mycoplasma gallisepticum.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 171..2156  
 FT /\*tag= a  
 FT misc\_difference 828..830  
 FT /\*tag= b  
 FT misc\_difference 978..980  
 FT /\*tag= c  
 FT misc\_difference 1083..1085  
 FT /\*tag= d  
 FT misc\_difference 1911..1913  
 FT /\*tag= e  
 FT /\*codon= seq: TGA, aa: Trp  
 PN JP07236498-A.  
 PD 12-SEP-1995.  
 XX 25-FEB-1994; 94JP-0052764.  
 PF 25-FEB-1994; 94JP-0052764.  
 PR 25-FEB-1994; 94JP-0052764.  
 XX  
 PA (JAPC ) NIPPON ZEON KK.  
 PA (SHIO ) SHIONOGI & CO LTD.  
 DR MPI; 1995-347462/45.  
 DR P-PSDB; AAR79910.  
 XX  
 PT Detection of Mycoplasma gallisepticum - for the quick detection,  
 PT i.e. within one day, of avian chronic respiratory diseases  
 XX  
 PS Claim 2; Page 7-10; 11pp; Japanese.  
 XX  
 XX This is the nucleotide sequence of a fragment of the Mycoplasma  
 CC gallisepticum genome which codes for a 661 amino acid protein.  
 CC This sequence and the sequence of AAT04076 (encoding a 261 amino  
 CC acid protein) can be used to detect M.gallisepticum using probes  
 CC based on nucleotides 1125-1648 and primers based on nucleotides  
 CC 449-466, the complement of bases 893-919, 1908-1934 and the  
 CC complement of bases 2184-2210 of this sequence, and a probe based  
 CC on nucleotides 718-41 of AAT04076. The method using these sequences  
 CC is faster i.e. is able to detect M.gallisepticum, which causes avian  
 CC chronic respiratory diseases, within one day, from avian secretions,  
 CC washings from the lung, respiratory tract, nasal cavity, etc.  
 XX  
 SQ Sequence 2196 BP; 800 A; 353 C; 376 G; 667 T; 0 other;

QY Query Match 16.6%; Score 216.6; DB 16; Length 2196;  
 Best Local Similarity 55.1%; Pred. No. 1.5e-26;  
 Matches 548; Conservative 0; Mismatches 414; Indels 33; Gaps 5;  
 QY 329 AATTAGAACGAGCGGATGAGTTAAGATCTAATCATGCTAAGAGCATGATTAAG 388

D	352	AATTGACAGACGCAAAAGCTGCTTTA	CTACTTGTGATATGCTGAACCTGC	AAATCTTG	411	
Q	359	CTTCACTACAAAGCTATGCCAAAGAT	TGAAGCTAGTTATCATCTGCTTAT	AGTGAAGCTG	448	
D	412	CGTCATATGMAACATATGCTTAAGAT	CAAAAGTAAATTAATCATCAGCGT	ATGAAACAGCTA	471	
Q	449	AAACAGTTAACATTAACCTTAATG	CAACATTTAGAACACATAAATGGC	CTAAACATAT	508	
D	412	AAGCAGTTCACTTAAACCTGCTG	CACTCTTAATAGGTTAATGAGC	AAAACTACAT	531	
Q	509	TAGAAATCAGCCATCAACCAAGCT	ATATCGGATTAACGCTTTGATTA	TGAAACACCCAA	568	
D	532	TAGATGCTGCTTATTAATAAAGCT	GTAGTACGTAAGATGATTTTGAT	GCACACAGGGT	591	
Q	569	ATTAGTTGAACCATCAAGAAGCT	TAACCACTTAAAGCAACGTCAT	ACCTTG	628	
D	552	CACATGCTGMAACATATTAACA	CTTAAGAAACGTTAAAGAAACAA	AACTAATTTAG	651	
Q	629	AAGCTTTGTCATCAACCTGCTT	ATATCAAAATTCGACATATTTAG	TGAGATCTATACATA	688	
D	652	ATTCTCTTCCAAACGAAATTTAT	GACGACATGAACTATCTTAAT	AGTTTATGAAA	711	
Q	669	AAGCTAGTATGTTAATTAATCA	AAACCTATGATCTAAATGGG	GGACGCTTTTGAT	748	
D	712	AAGCCAAATACATATGTTAC	AGCTATCTTGACCCCTGCTCTG	GAATAATCTCCGAAGTTA	771	
Q	749	CTAATGATGATCTACAGTTAAT	CGGAATTTATATATAGTATCA	ACTATTAATGAAC	808	
D	772	TGAGTG---TAA	CACAGCTAATTCAGATATTTCT	ATATGCACTTAATAGCTT	828	
Q	809	AAAAGACTAATGCTGATG	CAATTAATAGTTTATTTAAAAA	AGGATTCAAATATATG	868	
D	829	GAATAAATAATCTGATTAAT	TAGCTTAACAGTTTATCAAC	AGCTCTTTAGTTAAAAATA	888	
Q	869	AACAAGTTTGTAGGGA	CTTTTACAAAGCTAATTTTCA	ACCTTCAAACTACAGTTTG	928	
D	889	ATTTAATCTAGAGTTG---	ATGATGACAAATATATCAG	ACACGCAACGAAATTAACAGTTTG	945	
Q	929	TTGCTTTAGTGTGATGTA	CAACCCGCTCAATTAATTAAT	ATGACMAAGAGCCGTTTGA	988	
D	946	TTGGTTTATGTTATATGTT	ATGTAATCTTAATGAAATTTTG	CGCAAGAAAGAAAGTTTGGG	1000	
Q	969	ATGATGATGAACCTTCA	GTAGATTTCTTGCAACCGAAT	TAGTA-----	1033	
D	1006	CCTGTGAATAATCTCTTT	AGCACTACACAGCTGAAGT	GCACAAACAACAGCTGCAT	1066	
Q	1034	---TACAGATTTTCTTG	ATTTATAGTTAGCTGAA	CAACAACGAAGTACCAATTTTA	1099	
D	1066	CCTTAACAGATTTTCA	TGATCTAATAGTTAATGTG	CTGAAGCTAATACACTTA	1112	
Q	1091	GTTTATGCACTATG---	TCATCAACGTGTTATTTAT	TTTCCCTTAATAGTTGTTA	1114	
D	1126	GCTTTCGTACTTGTGA	AGCTGAATAAAGCTTACTTA	TATTTCCCTTAATTAATTTAGTTA	1118	
Q	1148	AAGCAGCTGATCTTA	ATAGTTGATTAACAATATACAA	TATTAATATGAAATGTTCAAC	1200	
D	1186	AAAC-----TAG	ATATATGTTTACATATATAG	TTTAATGTTAGTGTGATATCTAAC	1233	
Q	1208	AAATTGAGTTTCCACTT	CAACTGATGCAATATATATCA	AGCTTAATCCAACTCCACAG	1266	
D	1240	AAATTAACTTTGTAC	AAATCCACAGCTTCTG	GTTCAGATGATGTTGCTG	TATGAAGAAAG	1299
Q	1268	TTGATGAGATTAAG	TGTTGCTTAATATGCTTTAT	ATCA	1302	
D	1300	AAACTATGCTATGCT	CTGTAATAATGACAGTCA	1334		

```

RESULT 14
AAQ77853
ID      AAQ77853 standard; DNA; 2369 BP
XX
AC      AAQ77853;

```

XX	25-MAR-2003	(updated)	
DT	23-JUN-1995	(first entry)	
XX			
DE	Mycoplasma gallisepticum antigen TM-81 coding sequence.		
XX			
KW	recombinant avipox virus; live vaccine; mycoplasma antigen; ds.		
XX			
OS	Mycoplasma gallisepticum.		
XX			
PH	Key	Location/Qualifiers	
XX	CDS	171..2156	
FT		/*tag= a	
FT	mat_peptide	/product= antigenic_polypeptide	
FT		171..2153	
XX		/*tag= b	
XX		/codon= seq:TGA,aa:Trp	
PN	WO9423019-A1.		
PD	13-OCT-1994.		
XX			
PF	31-MAR-1994;	94WO-JP00541.	
XX			
PR	31-MAR-1993;	93JP-0074139.	
PR	30-SEP-1993;	93JP-0245625.	
XX			
PA	(JAP) NIPPON ZEON KK.		
PA	(SHIO ) SHIONOGI & CO LTD.		
XX			
PI	Aoyama S, Funato H, Iritani Y, Ohkawa S, Ohsawa I;		
PI	Saeki S, Saito, Takahashi K;		
XX			
DR	WPI; 1994-333181/41.		
XX	P-PSDB; AAR63226.		
PT	Recombinant avipox virus combining DNA encoding a polypeptide -		
PT	exhibiting antigenicity of mycoplasma, useful for the production		
XX	of a live vaccine		
PS	Claim 4; Page 61-65; 123pp; Japanese.		
XX			
CC	A restriction fragment of the insert of M.gallisepticum genomic clone		
CC	pUM-81 containing an open reading frame was sequenced (AAQ77853). The		
CC	ORF encodes an antigenic polypeptide. A recombinant avipox virus		
CC	comprising the coding sequence can be used as a live vaccine to		
CC	protect against infection by Mycoplasma gallisepticum.		
CC	(Updated on 25-MAR-2003 to correct FN field.)		
XX			
SO	Sequence 2369 BP; 859 A; 376 C; 401 G; 733 T; 0 other;		
	Query Match	16.6%; Score 216.6; DB 15; Length 2369;	
	Best Local Similarity	55.1%; Pred. No. 1.5e-26;	
	Matches 548; Conservative 0; Mismatches 414; Indels 33; Gaps 5		
QY	329 AATTAGAAGCAGCCGAGATGAGATTAAACAGATCTAATCAATGCTAAAGCGATGACATTAG 388		
Db	352 AATTAGCAGAGCAAAAGCTGCTTAATCTAATTGATTAAATGGTAAAGTGCATTCATG 411		
QY	389 CTTCACTACAAAGACTATGCCAAGATTGAAGCTAGTTATCATCTGCTTATAGTGAAGCTG 448		
Db	412 CGTCAATATGAAAGACTATGCTAAAGATCAAAAGTGAATTAAACATCAGCGATGAAACAGCTA 471		
QY	449 AAAAGATTAAACAATTAACCTTTATGCCAATATAGAACAACTAAATATGAGCTAAACTAATT 508		
Db	472 AAGCAGTTTACGCTTAAACCTGTGCAATCTTAATGAGTTAAATGAGGCAAAATCTACAT 531		
QY	509 TAGAATCAGCCATCAACCAAGCTAATACGATTAATAACGACTTTGATATGAACACCCAA 568		
Db	532 TAGATGCTGCTATTTAAAAAGCTGCTGATGCTAAGATGATTTTGATGACACAGCAGGCT 591		
QY	569 ATTATGTTGAAGCATACAAAGCACTAAACCACTTTAGAACAAAGTGTCTATCTAACCCTTG 628		

```

Db 592 CACTAGTGAAGCATATTAACATCTAAAGAAAGCTTAAAGAGAAAACTAATTAG 651
QY 629 AAGTTTGTCATCACTGCTTATATCAAAATTCGCAATATTTAGGATCTATACAA 688
Db 652 ATTCTCTGCAACGAAAAATTATGACGACATCGAACTAATCTTAATAGTTATAGAA 711
QY 689 AAGCTAGTATTAATACTAAACACTAGATCCACTAATATGGGGAGACGCTTTGAT 748
Db 712 AGCCAACTACTATTTGACGCTCTTACGCTTACCTCTCTCTGAAATATCTCGAAGTTA 771
QY 749 CTATAGATTAATACTAAGTTAATCGAATATTAATAATAGTTATCACTAATTAATGA 808
Db 772 TCAAGTATTAACACAGCTAATCAAGATATTACTAATCACTCAAGCTTAATAGCTT 828
QY 809 AAAAGCTAATGCTGATGATTAATCTAATAGTTTATTAAGAAAGATCTCAAAATATG 868
Db 829 GAAACAAATATGCTAATATTAATAGTAAACAGTTTATCAACAGCTTTAGTTAAATA 888
QY 869 AACAAAGTTTGTAGGACCTTTTACAAAGCTAATATGCACTTCAACTACAGTTTG 928
Db 889 ATTGACTAGAGTTG---ATGACGAAATATACAGAGAACACGAAATTTACAGTTTG 945
QY 929 TTGCTTTAGTGTGATGTATACACCGTCAATTAATATATGCAAGAGAACGTTTGA 988
Db 946 TTGTTTATGTTATATGTTGATCTCTAAGCTAAATTTTGGCAAGAAAGTTTGG 1005
QY 989 ATGGATTAACCTTCAAGTAAATTTCTGCAACAGCAATGTA----- 1033
Db 1006 CCTGAAATATCTCTTTAGCAACACACGCTGAAAGATCAACACAAAGCTGCAT 1065
QY 1034 ---TCACATGTTTCTGATTTATAGTTTACGCAACAAACAGAGTACCAATTA 1090
Db 1066 CCTTAACAGATTTTCAATATCTAATAGTTTAAATGCTGACGTAATACATTA 1125
QY 1091 GTTTAGCACTATG---TCATCACTGTTATTAATTTCCCTTAATAGTTGTTA 1147
Db 1126 GCTTTCGTTACTTTGAGCTGAAAAACAGCTTACTTATTTCCCTTAATAGTTA 1185
QY 1148 AAGCAGCTGATCTAATACGTTGATTCATCACTAATTAATATGAATGTTCAAC 1207
Db 1186 AAAC-----TAGATTAATGTTGTTTACAAATATTAATGATGATTAAC 1239
QY 1208 AAGTGAATTCACCTCACTAGTCAATATTAATCTACAGCTAATCACTCCAGCAG 1267
Db 1240 AATTAATCTTTTACAACTCCAGCTTCTGTTCAAGTATGTTCTGCTATGAAGA 1299
QY 1268 TTGATGAGTTAAAGTTGCTAAATCGTTTATCA 1302
Db 1300 AAATATGCTATGCTCTGCTGAAATGACAGTCAGCA 1334

RESULT 15
AA068670 standard; DNA; 2452 BP.
XX
AC AA068670;
XX
DT 28-FEB-1995 (first entry)
XX
DE PMGAI.3 Mycoplasma gallisepticum DNA.
XX
KW PMGAI.3 Mycoplasma gallisepticum; conserved sequences;
XX primer; probes; amplification; polymerase chain reaction;
XX specific; detection; PCR; 1.2; 1.3; ss.
XX
OS Mycoplasma gallisepticum.
XX
PN AU930593-A.
XX
PD 26-MAY-1994.
XX
PF 10-NOV-1993; 93AU-0050593.
XX

```

```

PR 10-NOV-1992; 92AU-0005744.
XX
XX (UYME) UNIV MELBOURNE.
PA (BROW/) BROWNING G F.
XX
XX Browning GF, Glew MD, Markham PF, Walker ID, Whitehead KG;
PI WPI; 1994-209061/26.
XX
XX Recombinant DNA constructs for Mycoplasma gallisepticum - for
PT diagnosis, treatment and prophylaxis of poultry respiratory
XX disorders.
XX
XX Example 1; Fig 4; 51bp; English.
XX
XX AA068670 shows the DNA sequence of PMGAI.3 derived from a 10 kb insert
XX from a Mycoplasma gallisepticum genomic DNA library, detected by probes
XX based on the T3 and C7 peptides. PMGAI.3 contains a (GAA)n consensus
XX sequence (n=4 to 13). Mycoplasma gallisepticum infection in poultry,
XX humans and other animals is of economic importance to many industries
XX and it is desirable to produce effective vaccines and probes for its
XX detection. The sequences and probes and vaccine vectors of the
XX infection can be used for the diagnosis and treatment of Mycoplasma
XX gallisepticum infection, and for prophylaxis.
XX
SQ Sequence 2452 BP; 913 A; 404 C; 420 G; 715 T; 0 other;

Query Match 15.7%; Score 205.6; DB 15; Length 2452;
Best Local Similarity 53.5%; Pred. No. 8.8e-25;
Matches 492; Conservative 0; Mismatches 404; Indels 24; Gaps 2;

QY 323 AAACCAATTAAGGACGCGGATGAGTTAACAAGATCTAATCAATGCTAAAGCATGA 382
Db 410 AACCAATTAAGGACGCGGATGAGTTAACAAGATCTAATCAATGCTAAGCAATACCA 469
QY 383 CATTAGCTTCACTCAAGACTATGCAAGATTTGAAGTATGATATCATCTGCTTAATAGT 442
Db 470 ATGTTCCTTTATGCTGATTAATGCAAAATTTCAAGATCTTTAAGCAGCTTATAGA 529
QY 443 AAGTGAACGTTAACAATTAACCTTAATGCAATTAAGCAATTAATTAATGCTTAA 502
Db 530 CAGCTAAACGCTATCAAGAAATTAACAAGCGCACTTTAAGAAATCTAAGATCTGCATCA 589
QY 503 CTAATTTAGATTCAGCATCAACAGCTAATAGGATTAATTAAGCACTTTGATATGAC 562
Db 590 CTACACTCAAGCAGCTATGATTAAGCTGCTAATGATTAAGCTGATTAATGATGTTA 649
QY 563 ACCCAATTTAGTGAAGCATACAAAGCATTAATAACCACTTGAACAACAGTCTACTA 622
Db 650 ATCAACCTTTAGTGAAGCATACAAAGCATTAATAACCTTGAACAACAGTCTACTA 709
QY 623 ACCTGAAGTTTGTATCAACGCTTATTAATCAATTTGCAATTAATTTAGTATCTAT 682
Db 710 GTTTGAAGGATTTATGCAAAATTAATGATGCTATTAATAATCACTAATGATCTGT 769
QY 683 ACAATTAAGCTAGTATTTAATACTAATAACATGATCCACTAATGCGGAGACGCTT 742
Db 770 TTGATACGAGTATGCAATTAACACTTAATAACATTAATGATCTAATGATGATGAACCTA 829
QY 743 TAGATCTAATGATTAATCACTAATGATTAATTAATTAATTAATCACTAATTA 802
Db 830 CTTTAGGAAAGTGAAGCAAGCTAACAACGCTATTAATAATGCTATTAATGCTAATGCT 889
QY 803 ATGAACAAAGACTAATGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTA 862
Db 890 TGAAGAAATGAAGCAAGCAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTA 949
QY 863 ATATGAACAAAGTTTGTAGGACCTTTCAAAACGCTAATGTTCAACCTTCAACACTACA 922
Db 950 AAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1009
QY 923 GTTTGTTGCTTTAGTGTATGATTAAC-----ACCGCTAATTAATTAAT 967

```

```

Db      1010 GTTTCAGCTTAAGTGTGATCTAAGTAAATTCGCAAAATCTAAGTAAAT 1069
Qy      968 ATGCAAGAAGGACCGTTGGAATGGTGAATGAACCTTCAAGTAGAA-----TTCCTTG 1018
Db      1070 TTGCTCAAGAAGGATTGAACTAGTGAATAATCAACAAACCGGAAAAACAGCTAGTTT 1129
Qy      1019 CAACACGAATAGTATCAACAGATGTTTCTTGATTTATAGTTTACCTGGAACAAACACGA 1078
Db      1130 CCTCTCCTGTTTCTGCAACTGATGTTTCTTGATTTATAGTTTACCTGCTGGAAGGACTA 1189
Qy      1079 AGTACCAATTTAGTTTACCACTATGTCCTCACTGCTTATTTATTTCCCTATA 1138
Db      1190 AATATACATTAACCTTGAATTAATGATCCGATTAATGCAATCTTATATTTGCCCTATA 1249
Qy      1139 AGTGGTTAAAGCAGCTGATGCTAATAAGTTGATTAACAATACAAATTAATAATGGA 1198
Db      1250 AGTGGTTAAAGCAGCTGATTAATCAAGTTCAGTAGCCCTTCAATACAGCTTAATAAATA 1309
Qy      1199 ATGTTCAACAAGTTGAGTTT 1218
Db      1310 GTTCGAATTAATTAATTTT 1329

```

Search completed: August 25, 2003, 03:27:35  
 Job time : 423 secs

THIS PAGE BLANK (USPS)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 25, 2003, 03:15:12 ; Search time 98 Seconds  
(without alignments)  
5882.106 Million cell updates/sec

Title: US-09-901-572A-1  
Perfect score: 106  
Sequence: 1 aaaaacatcgatcttcaat.....taaacgtttatcaggtt 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1280.4	98.0	2144	2	US-08-525-742-9
2	1270	97.2	1387	2	US-08-525-742-3
3	1264	96.8	1305	1	US-08-185-851A-1
4	1086.8	83.2	2014	2	US-08-525-742-7
5	216.6	16.6	2369	2	US-08-525-742-1
6	65.6	5.0	3057	4	US-09-601-198-55
7	55	4.2	10640	4	US-09-417-485D-5
8	53.6	4.1	1887	4	US-09-601-198-39
9	53.2	4.0	8700	2	US-08-645-193B-18
10	51.6	4.0	8700	2	US-08-392-625-16
11	51.6	4.0	8700	2	US-08-466-961A-16
12	51.6	4.0	53332	4	US-09-801-861-3
13	50.8	3.9	1851	4	US-09-601-198-51
14	50.6	3.9	14066	4	US-09-601-198-56
15	50	3.8	2043	4	US-09-601-198-181
16	49.4	3.8	615	3	US-08-998-416-186
17	49.2	3.8	837	3	US-08-998-416-288
18	49.2	3.8	8920	2	US-08-446-855A-1
19	49.2	3.8	8920	3	US-09-150-741-1
20	48.6	3.7	636	3	US-08-998-416-1137
21	48.2	3.7	1500	4	US-09-601-198-36
22	48.2	3.7	19124	2	US-08-487-826B-13
23	48	3.7	19124	2	US-08-487-826B-13
24	47.6	3.6	827	3	US-08-998-416-535
25	46.8	3.6	732	3	US-08-998-416-1036
26	46.6	3.6	2427	4	US-09-601-198-70
27	46.4	3.6	20674	4	US-09-641-638-651

28	46.2	3.5	701	3	US-08-998-416-701	Sequence 701, App
29	46.2	3.5	15016	4	US-09-601-198-60	Sequence 60, Appl
30	46	3.5	998	3	US-09-122-400B-5	Sequence 5, Appl
31	45.8	3.5	1431	3	US-09-316-083-2	Sequence 2, Appl
32	45.8	3.5	1431	4	US-09-933-700-2	Sequence 1, Appl
33	45.8	3.5	1830121	4	US-09-557-884-1	Sequence 1, Appl
34	45.8	3.5	1830121	4	US-09-643-990A-1	Sequence 1, Appl
35	45.6	3.5	3399	4	US-09-601-198-43	Sequence 43, Appl
36	45.6	3.5	6124	4	US-08-213-419B-3	Sequence 3, Appl
37	45.4	3.5	945	4	US-09-601-198-177	Sequence 177, App
38	45.4	3.5	3509	4	US-09-255-829-19	Sequence 19, Appl
39	45.4	3.5	202001	4	US-09-734-674-3	Sequence 3, Appl
40	45	3.4	5340	4	US-09-627-122-21	Sequence 21, Appl
41	44.8	3.4	30549	4	US-09-134-001C-322	Sequence 322, App
42	44.6	3.4	6243	2	US-09-056-075-1	Sequence 1, Appl
43	44.6	3.4	6265	4	US-09-129-112-3	Sequence 3, Appl
44	44.6	3.4	1664976	4	US-08-916-421B-1	Sequence 1, Appl
45	44.4	3.4	658	3	US-08-998-416-595	Sequence 595, App

## ALIGNMENTS

RESULT 1  
US-08-525-742-9  
; Sequence 9, Application US/08525742  
; Patent No. 5871742  
; GENERAL INFORMATION:  
; APPLICANT: Saito, Shuji  
; APPLICANT: Ohkawa, Setsuko  
; APPLICANT: Sasaki, Sakiko  
; APPLICANT: Ohnawa, Ikuroh  
; APPLICANT: Funato, Hiroko  
; APPLICANT: Iritani, Yoshikazu  
; APPLICANT: Aoyama, Shigema  
; APPLICANT: Takahashi, Kiyoohito  
; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
; TITLE OF INVENTION: AS USE THEROF  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
; ADDRESSEE: NAUGHTON  
; STREET: 1725 K Street, Suite 1000  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525, 742  
; FILING DATE: 25-SEP-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 05-074139  
; FILING DATE: 31-MAR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 05-245625  
; FILING DATE: 30-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP94/00541  
; FILING DATE: 31-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mcleland, Ie-Mhung  
; REGISTRATION NUMBER: 31,541  
; REFERENCE/DOCKET NUMBER: 950811  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-659-2930

```

? TELEPFX: 202-8870350
? INFORMATION FOR SEQ ID NO: 9
? SOURCE: CHARACTERISTICS:
? LENGTH: 2144 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 202..2046
?
US-08-525-742-9

```

Query Match	98.0%;	Score 1280.4;	DB 2;	Length 2144;
Best Local Similarity	98.8%;	Pred. No. 1.2e-258;		
Matches 1290;	Conservative 0;	Mismatches 16;	Indels 0;	Gaps 0

OY	1	AAAAACATCAGATGGTAAATCTGATATCTTGGCTTAAAAAAAACAACAAATCTCTAACAA	60
Db	1	AAAAACATCAGATGGTAAATCTGATATCTTGGCTTAAAAAAAACAACAAATCTCTAACAA	60
OY	61	AATCTAAATTAATAAGCCGTTAAATTAACCTAAAAAATTTAAAAAATGTTTTCTTATC	120
Db	61	AATCTAAATTAATAAGCCGTTAAATTAACCTAAAAAATTTAAAAAATGTTTTCTTATC	120
OY	121	AACCAAAATTCCTAGTAATTAACCGCTTATTTATTTTATTTAGCATCTTTAAGAT	180
Db	121	AACCAAAATTCCTAGTAATTAACCGCTTATTTATTTTATTTAGCATCTTTAAGAT	180
OY	181	ATPAATATATCTTAATTAATTCATGATTAAGAAAAGAAATCATCTTAAAGCATATAGTTG	240
Db	181	ATPAATATATCTTAATTAATTCATGATTAAGAAAAGAAATCATCTTAAAGCATATAGTTG	240
OY	241	TTAGGTAACAATCCTTTCTTAGCATTTGGATTTTCTAGCTGATGCTATTAACAAAAA	300
Db	241	TTAGGTAACAATCCTTTCTTAGCATTTGGATTTTCTAGCTGATGCTATTAACAAAAA	300
OY	301	GATGCAAAACCCAAATTAATGGCCAAACCCAAATTGAAGCAGCGGGATGAGTTAACAGAT	360
Db	301	GATGCAAAACCCAAATTAATGGCCAAACCCAAATTGAAGCAGCGGGATGAGTTAACAGAT	360
OY	361	CTAATCAATGCTTAAGCGATGACATTAGCTTCACTACAAGCATCTGCCAAGATTGAAGCT	420
Db	361	CTAATCAATGCTTAAGCGATGACATTAGCTTCACTACAAGCATCTGCCAAGATTGAAGCT	420
OY	421	AGTTATCATCTGCTTATATAGTGAAGCTGAACGTTAACAATTAACCTTAATGCAATTA	480
Db	421	AGTTATCATCTGCTTATATAGTGAAGCTGAACGTTAACAATTAACCTTAATGCAATTA	480
OY	481	GAAACAATTAATGGCTAAAACTAATTTTGAATCAAGCATCAACCAAGTATTAAGGAT	540
Db	481	GAAACAATTAATGGCTAAAACTAATTTTGAATCAAGCATCAACCAAGTATTAAGGAT	540
OY	541	AAAAAGACTTTGATPATGAACAACCCAAATTTAGTTGAAGCATCAAGACATTAANAAC	600
Db	541	AAAAAGACTTTGATPATGAACAACCCAAATTTAGTTGAAGCATCAAGACATTAANAAC	600
OY	601	ACTTTAGAACACGCTGCTACTAACCTTGAAGTTTGTCATCAACCTGCTTAATGCAAT	660
Db	601	ACTTTAGAACACGCTGCTACTAACCTTGAAGTTTGTCATCAACCTGCTTAATGCAAT	660
OY	661	CGCAATTAATTTAGTGATCTATACAAATAAGCTAGTATTAATTAACCTAAACACATAGAT	720
Db	661	CGTATTAATTTAGTGATCTATACAAATATGCTAGTAGTTTANTATTAACCTAAACACATAGAT	720
OY	721	CCACTAAATGSGGGAACGCTTTTGAATTCCTAATGAGATTACTACAGTTATCGAATATT	780
Db	721	CCACTAAATGSGGGAATGCTTTTGAATTCCTAATGAGATTACTACAGTTATCGAATATT	780
OY	781	AATATAGGTTATGCACTATTAATGAACAAAAGACTAAAGCTGATGATTAATCTAATAGT	840
Db	781	AATATAGGTTATGCACTATTAATGAACAAAAGACTAAAGCTGATGATTAATCTAATAGT	840

QY	841	TTTATTAATAAAGGCGATTCAAATATATAGAACAAAGTTTGATGGGACTTTTACAAACGCT	900
Db	841	TTTATTAATAAAGGCGATTCAAATATATAGAACAAAGTTTGATGGGACTTTTACAAACGCT	900
QY	901	AATGTTCAACCTTCMAACTACAGTTTGTGTCTTTAGTGCATGATTAACACCCGTCAT	960
Db	901	AATGTTCAACCTTCMAACTACAGTTTGTGTCTTTAGTGCATGATTAACACCCGTCAT	960
QY	961	TATTAATATGCAAGAGGACCGTTTGGAAATGCTGATGAACCTTCAAGTGAATTTCTTGA	102
Db	961	TATTAATATGCAAGAGGACCGTTTGGAAATGCTGATGAACCTTCAAGTGAATTTCTTGA	102
QY	1021	AACAGATAGTATCACAGATGTTTCTTGGATTTTATAGTTTAGCTGGAACAAACAGAG	108
Db	1021	AACAGATAGTATCACAGATGTTTCTTGGATTTTATAGTTTAGCTGGAACAAACAGAG	108
QY	1081	TACCAATTTAGTTTATGCACTAGTGTCCATCAACTGTTTATTTATTTCCCTTATAAG	114
Db	1081	TACCAATTTAGTTTATGCACTAGTGTCCATCAACTGTTTATTTATTTCCCTTATAAG	114
QY	1141	TTGGTTAAGAGCGATGCTATATACGTTGGATTATACATACAAATTAATATAGGAAT	120
Db	1141	TTGGTTAAGAGCGATGCTATATACGTTGGATTATACATACAAATTAATATAGGAAT	120
QY	1201	GTTCAACAGTTGAGTTTGCCACTTCAACTAGTGCMAATATATCTACAGCTATCCAACT	126
Db	1201	GTTCAACAGTTGAGTTTGCCACTTCAACTAGTGCMAATATATCTACAGCTATCCAACT	126
QY	1261	CCACGAGTTAGTGATTTAAAGTGTCTAAATTCGTTTATACAGGTT	1306
Db	1261	CCACGAGTTAGTGATTTAAAGTGTCTAAATTCGTTTATACAGGTT	1306

RESULT 2  
 US-08-525-742-3  
 Sequence 3, Application US/08525742  
 Patent No. 5871742  
 GENERAL INFORMATION:  
 APPLICANT: Saito, Shuji  
 APPLICANT: Ohkawa, Setsuko  
 APPLICANT: Saeki, Sakiko  
 APPLICANT: Ohnawa, Ikuroh  
 APPLICANT: Funato, Hiroko  
 APPLICANT: Iritani, Yoshikazu  
 APPLICANT: Aoyama, Shigemi  
 APPLICANT: Takahashi, Kiyoochito  
 TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
 TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
 TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
 NUMBER OF SEQUENCES: 51  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, MCLELAND &  
 ADDRESSEE: NAUGHTON  
 STREET: 1725 K Street, Suite 1000  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/525,742  
 FILING DATE: 25-SEP-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 05-074139  
 FILING DATE: 31-MAR-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 05-245625

FILING DATE: 30-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/JP94/00541  
 FILING DATE: 31-MAR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mclelland, Le-Nhung  
 REGISTRATION NUMBER: 31,541  
 REFERENCE/DOCKET NUMBER: 950811  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-659-2930  
 TELEFAX: 202-8870357  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1387 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 202..1305  
 US-08-525-742-3

Query Match 97.2%; Score 1270; DB 2; Length 1387;  
 Best Local Similarity 98.8%; Pred. No. 1.6e-256;  
 Matches 1290; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

1 AAAAAACAGATGTTGATCTGATATCTTGTAAACAAACAAATCTTCAACAA 60  
 1 AAAAAACAGATGTTGATCTGATATCTTGTAAACAAACAAATCTTCAACAA 60  
 61 AATCTTAATTAATTAAGCCGTTAAATTAATAAATAAATGTTTCTTATC 120  
 61 AATCTTAATTAATTAAGCCGTTAAATTAATAAATAAATGTTTCTTATC 120  
 121 AACCAAAATCTCTAGTAATAAAGCGTTATTTTATTTTATTTAGTCACTTTTAAAGAT 180  
 121 AACCAAAATCTCTAGTAATAAAGCGTTATTTTATTTTATTTAGTCACTTTTAAAGAT 180  
 181 AATAAATATCTTATATATTTCTATGATGAAGAAGAAATCACTTAAACATTTAGTTG 240  
 181 AATAAATATCTTATATATTTCTATGATGAAGAAGAAATCACTTAAACATTTAGTTG 240  
 241 TTAGGTACACATCTTCTTAGCATGAGATTTCTAGCTGATGCTATTACTTAAACAA 300  
 241 TTAGGTACACATCTTCTTAGCATGAGATTTCTAGCTGATGCTATTACTTAAACAA 300  
 301 GATGCAAAACCAAAATTAATGCGCAAAACCAATTTAGAGCAGCGCAATGAGTTAAACAGAT 360  
 301 GATGCAAAACCAAAATTAATGCGCAAAACCAATTTAGAGCAGCGCAATGAGTTAAACAGAT 360  
 361 CTAATCAATGCTAAAGCATGACATTAGCTTCACTCAAGCACTATGCCAAGATTGAAGCT 420  
 361 CTAATCAATGCTAAAGCATGACATTAGCTTCACTCAAGCACTATGCCAAGATTGAAGCT 420  
 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACTAATTAATGCAACATTA 480  
 421 AGTTATCATCTGCTTATAGTGAAGCTGAACAGTTAACTAATTAATGCAACATTA 480  
 481 GAACAACCTAAATGCTTAAATCTAATTTAGATCAGCATCAACCAAGCTTAATCGAGAT 540  
 481 GAACAACCTAAATGCTTAAATCTAATTTAGATCAGCATCAACCAAGCTTAATCGAGAT 540  
 541 AAAACGACTTTTGAATTAAGAACCAATTTAGTGAAGCATTAACCAAGCTTAATCGAGAT 600  
 541 AAAACGACTTTTGAATTAAGAACCAATTTAGTGAAGCATTAACCAAGCTTAATCGAGAT 600  
 601 ACTTTGAACAACGCTGCTACTTAATGAGGTTTGTATCACTCACTTATTAATCAAT 660  
 601 ACTTTGAACAACGCTGCTACTTAATGAGGTTTGTATCACTCACTTATTAATCAAT 660  
 661 CGCAATTAATTTAGTGAATCTATACAAATAAGCTAGTATTTAATTAATTAACCTAGAT 720  
 661 CGCAATTAATTTAGTGAATCTATACAAATAAGCTAGTATTTAATTAATTAACCTAGAT 720

DB 661 CGTAAATATTTAGTGAATCTATACAAATAAGCTAGTATTTAATTAATTAACCTAGAT 720  
 QY 721 CCCTAAATATGAGGAGAGCTTTTATGATTTCTAATGAGATTACTACAGTTAATCGAATAT 780  
 DB 721 CCCTAAATATGAGGAGAGCTTTTATGATTTCTAATGAGATTACTACAGTTAATCGAATAT 780  
 QY 781 AATAATAGTTATCACTATTAATGAACAAAGACTAATGCTGATGCTATTAATAGT 840  
 DB 781 AATAATAGTTATCACTATTAATGAACAAAGACTAATGCTGATGCTATTAATAGT 840  
 QY 841 TTTATTAATAAAGAGATTCAAATAATGAACAAAGCTTTTATGAGGACTTTTAAACAGCT 900  
 DB 841 TTTATTAATAAAGAGATTCAAATAATGAACAAAGCTTTTATGAGGACTTTTAAACAGCT 900  
 QY 901 AATGTTCACTTCAAACTAATGATGTTTGTGCTTTTATGCTGATGCTAATCAACCCGTCAT 960  
 DB 901 AATGTTCACTTCAAACTAATGATGTTTGTGCTTTTATGCTGATGCTAATCAACCCGTCAT 960  
 QY 961 TATAAATATGCAAGAGAGAGCCGTTTGAATGATGATGAACCTTCAAGTAAATCTTCA 1020  
 DB 961 TATAAATATGCAAGAGAGAGCCGTTTGAATGATGATGAACCTTCAAGTAAATCTTCA 1020  
 QY 1021 AACAGCAATGATCAAGATGTTTCTTGAATTTATGATTTAGTGAACCAACAGAG 1080  
 DB 1021 AACAGCAATGATCAAGATGTTTCTTGAATTTATGATTTAGTGAACCAACAGAG 1080  
 QY 1081 TACCAATTTAGTTTATGCAATGATGCTCAATGCTGATTTATTAATTTCCCTATTAAG 1140  
 DB 1081 TACCAATTTAGTTTATGCAATGATGCTCAATGCTGATTTATTAATTTCCCTATTAAG 1140  
 QY 1141 TTGTTTAAAGCAGCTGATGCTAATTAATGATTTATTAACAAATTAATTAATGAAAT 1200  
 DB 1141 TTGTTTAAAGCAGCTGATGCTAATTAATGATTTATTAACAAATTAATTAATGAAAT 1200  
 QY 1201 GTTCAACAGATTGATTTGCCACTTCACTATGATGCAATTAATTAATTAATGAAAT 1260  
 DB 1201 GTTCAACAGATTGATTTGCCACTTCACTATGATGCAATTAATTAATTAATGAAAT 1260  
 QY 1261 CCAGCAGTGTATGATTAAGTTGCTAAATGCTTTATGAGTT 1306  
 DB 1261 CCAGCAGTGTATGATTAAGTTGCTAAATGCTTTATGAGTT 1306

RESULT 3  
 US-08-185-851A-1  
 ; Sequence 1, Application US/08185851A  
 ; Patent No. 5489430  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saito, Shuji  
 ; APPLICANT: Ohkawa, Setsuko  
 ; APPLICANT: Fujisawa, Ayumi  
 ; APPLICANT: Iritani, Yoshikazu  
 ; APPLICANT: Aoyama, Shigemitsu  
 ; TITLE OF INVENTION: Poultry Mycoplasma Antigens, Gene  
 ; TITLE OF INVENTION: Therapeutic and Recombinant Vectors Containing the Gene As  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Armstrong, Westerman, Hattori, Mclelland &  
 ; ADDRESS: Naughton  
 ; STREET: 1725 K Street, N.W., Suite 1000  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS version 5.0  
 ; SOFTWARE: ASCII from Word Perfect version 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/185,851A  
 ; FILING DATE:

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Theresa M. Stevens-Smith
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: PO-8-A930918
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-887-0357
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 202..1308
US-08-185-851A-1

```

```

Query Match      96.8%; Score 1264; DB 1; Length 1305;
Best Local Similarity 98.3%; Pred. No. 2,8e-255;
Matches 1284; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

```

```

QY 1 AAAAAATCATGATTTGTTAACTGATATCTTTGTTAAACACAAATCTTCAACA 60
DB 1 AAAAAATCATGATTTGTTAACTGATATCTTTGTTAAACACAAATCTTCAACA 60
QY 61 AATCTTAATTAATTAAGCGTTAAATTAATACTAAATTAATTAATTAATTAAT 120
DB 61 AATCTTAATTAATTAAGCGTTAAATTAATACTAAATTAATTAATTAATTAAT 120
QY 121 AACCAAAATCTCTAGTAAATTAAGCTTTATTTATTTTATTTTATTTTATTTT 180
DB 121 AACCAAAATCTCTAGTAAATTAAGCTTTATTTATTTTATTTTATTTTATTTT 180
QY 181 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
DB 181 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 241 TTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
DB 241 TTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
QY 301 GATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
DB 301 GATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 360
QY 361 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
DB 361 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
QY 421 AGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
DB 421 AGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
QY 481 GAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
DB 481 GAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
QY 541 AAAAGACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
DB 541 AAAAGACTTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
QY 601 ACTTTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
DB 601 ACTTTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
QY 661 CGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
DB 661 CGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
QY 720 CGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
DB 720 CGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780

```

```

DB 721 CCACATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
QY 781 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
DB 781 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
QY 841 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
DB 841 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 901 AATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
DB 901 AATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
QY 961 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
DB 961 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
QY 1021 AACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
DB 1021 AACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
QY 1081 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
DB 1081 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
QY 1141 TTGTTAAAGAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
DB 1141 TTGTTAAAGAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
QY 1201 GTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
DB 1201 GTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1260
QY 1261 CCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1306
DB 1261 CCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1306

```

```

RESULT 4
US-08-525-742-7
Sequence 7, Application US/08525742
Patent No. 5871742
GENERAL INFORMATION:
APPLICANT: Saito, Shuji
APPLICANT: Ohkawa, Setsuko
APPLICANT: Saeki, Sakiko
APPLICANT: Ohsewa, Ikuroh
APPLICANT: Funato, Hiroyo
APPLICANT: Iritani, Yoshikazu
APPLICANT: Aoyama, Shigemi
APPLICANT: Takahashi, Kiyohito
TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE
TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND
TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESSES:
ADDRESSEE: ARMSTRONG, WESTERMAN, HAYTORI, MCLELAND &
ADDRESSEE: NAUGHTON
STREET: 1725 K Street, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742

```

FILING DATE: 25-SEP-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 05-074139  
 FILING DATE: 31-MAR-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 05-245625  
 FILING DATE: 30-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/JP94/00541  
 FILING DATE: 31-MAR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McLeand, Le-Nhung  
 REGISTRATION NUMBER: 31,541  
 REFERENCE/DOCKET NUMBER: 950811  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-659-2930  
 TELEFAX: 202-8870357  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2014 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 54..1883  
 US-08-525-742-7

Query Match 83.2%; Score 1086.8; DB 2; Length 2014;  
 Best Local Similarity 96.7%; Pred. No. 2,8e-218;  
 Matches 1120; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

150 TTTATTTTATTTTATTTAGT-CATCTTTAAGATATATATATCTTAATTTCTATGATA 208  
 1 TTTATTTTATTTTATTTTGGTAAATCTTTTAAATATATATATTTTAAATATTTCTATGATA 60  
 209 AGAAGAAGATCATCTTAAGACTATTAGTTTGGTAAACATCCCTTCTTATGACATG 268  
 61 AAAAAAGATCATCTTAAGACTATTAGTTTGGTAAACATCCCTTCTTATGATATG 120  
 269 GGATTTCTAGCTGTATGTCTATTAATAAAAAAGATGCAAAACCAATTAATGCCCC 328  
 121 GGATTTCTAGCTGTATGTCTATTAATAAAAAAGATGCAAAACCAATTAATGCCCC 180  
 329 AATTGAAGCAGCGGGAATGAGTTAACAAGATCTTAATCAATGCTAAAGCGATGACATTAG 388  
 181 AATTGAAGCAGCGGGAATGAGTTAACAAGATCTTAATCAATGCTAAAGCGATGACATTAG 240  
 389 CTTCACCTCAAGACATGCTCAAGATTTGAAGTATTCATTCGCTTATAGTGAAGCTG 448  
 241 CTTCACCTCAAGACATGCTCAAGATTTGAAGTATTCATTCGCTTATAGTGAAGCTG 300  
 449 AAACAGTTAACAATTAACCTTAATGCAACATTGAACAACCTAAATGCTAAATTT 508  
 301 AAACAGTTAACAATTAACCTTAATGCAACATTGAACAACCTAAATGCTAAATTT 360  
 509 TAGAATGAGCCATCAACCAAGCTAATACGATTAAGATTAAGATTAAGAACCCCA 568  
 361 TAGAATGAGCCATCAACCAAGCTAATACGATTAAGATTAAGATTAAGAACCCCA 420  
 569 AATTAGTTGAAGCACAAGACCTTAATGAACCTTTGAAACAGTCTCTAATCCTTG 628  
 421 AATTAGTTGAAGCACAAGACCTTAATGAACCTTTGAAACAGTCTCTAATCCTTG 480  
 629 AAGGTTTGCATCACTGCTTATATCAAAATTCGCAATTAATTTAGTGATCTATACATA 688  
 481 AAGGTTTGCATCACTGCTTATATCAAAATTCGCAATTAATTTAGTGATCTATACATA 540  
 689 AAGCTAGTAGTTTAATTAATAAACAAGCTAATGATCCATTAATGAGGGAACGCTTTAGATT 748  
 541 AAGCTAGTAGTTTAATTAATAAACAAGCTAATGATCCATTAATGAGGGAACGCTTTAGATT 600

749 CTAATGAGATTAACAGTTAATGGAATATTAATATAGCTTATCAACTATTATGAC 808  
 601 CTAATGAGATTAACAGTTAATGGAATATTAATATAGCTTATCAACTATTATGAC 660  
 809 AAAAGCTAATGCTGATGATCTTAATTAATGTTTATTAATAAAGATTCATAATATG 868  
 661 AAAAGCTAATGCTGATGATCTTAATTAATGTTTATTAATAAAGATTCATAATATATA 720  
 869 AACAAATTTTGGAGGACCTTTTAAACCGCTAATGTTCAACCTTCAACCTGATTTTG 928  
 721 AACAAATTTTGGAGGACCTTTTAAACCGCTAATGTTCAACCTTCAACCTGATTTTG 780  
 929 TTGCTTTAGTGTGATGATGATCAACCGCTCAATTAATTAATGCAAGAGACGTTTGA 988  
 781 TTGCTTTAGTGTGATGATGATCAACCGCTCAATTAATTAATGCAAGAGACGTTTGA 840  
 989 ATGGTATGAACCTTCAAGTGAATTTCTTCAACAGAAATGATGATCAAGATGTTTCTT 1048  
 841 ATGGTATGAACCTTCAAGTGAATTTCTTCAACAGAAATGATGATGATGATTTTCTT 900  
 1049 GGATTTATAGTTTATGCTGGAACCAACAGATTCATTTAGTTTATGCAATGATGCTC 1108  
 901 GAATTTATAGTTTATGCTGGAACCAACAGATTCATTTAGTTTATGCAATGATGCTC 960  
 1109 CATCACTGCTTATTTATTTTCCCTTATTAAGTTTGAAGCAGTGTATGATTAACG 1168  
 961 CATCACTGCTTATTTATTTTCCCTTATTAAGTTTGAAGCAGTGTATGATTAACG 1020  
 1169 TTGATTTACATTAACAATTAATTAATGAATGTTTCAACAGTTGATGATGATGATGAT 1228  
 1021 TTGATTTACATTAACAATTAATTAATGAATGTTTCAACAGTTGATGATGATGATGAT 1080  
 1229 CTAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1288  
 1081 CTAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140  
 1289 AAATGCTTTATGAGTT 1306  
 1141 AAATGCTTTATGAGTT 1158

RESULT 5  
 US-08-525-742-1  
 ; Sequence 1, Application US/08525742  
 ; Patent No. 5871742  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Saito, Shuji  
 ; APPLICANT: Ohkawa, Setsuko  
 ; APPLICANT: Saeki, Sakiko  
 ; APPLICANT: Ohnawa, Hiroyuki  
 ; APPLICANT: Funato, Hiroko  
 ; APPLICANT: Iritani, Yoshihiko  
 ; APPLICANT: Aoyama, Shigemi  
 ; APPLICANT: Takahashi, Kiyohito  
 ; TITLE OF INVENTION: NEW POLYPEPTIDE, DNA ENCODING THE  
 ; TITLE OF INVENTION: POLYPEPTIDE, RECOMBINANT VECTOR BEARING THE DNA AND  
 ; TITLE OF INVENTION: RECOMBINANT VIRUS UTILIZING THE RECOMBINANT VECTOR AS WELL  
 ; NUMBER OF SEQUENCES: 51  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ARMSTRONG, WESTERMAN, HATTORI, McLELAND &  
 ; STREET: 1725 K Street, Suite 1000  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/525,742
FILING DATE: 25-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-074139
FILING DATE: 31-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 05-245625
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP94/00541
FILING DATE: 31-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: McIceland, Le-Nhung
REGISTRATION NUMBER: 31,541
REFERENCE/DOCKET NUMBER: 950811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-659-2930
TELEFAX: 202-8870357
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2369 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 171..2153
US-08-525-742-1

```

Query Match 16.6%; Score 216.6; DB 2; Length 2369;

Best Local Similarity 55.1%; Pred. No. 1.1e-36; Matches 548; Conservative 0; Mismatches 414; Indels 33; Gaps 5;

```

339 AATTGAGACGAGCGGATGAGTGAATCAAGATCTAATCAATGCTTAAGCGATGACATTAG 388
352 AATTGAGACGAGCGGATGAGTGAATCAAGATCTAATCAATGCTTAAGCGATGACATTAG 411
389 CTTCCTACAGACGATGAGTGAATGAGTGAATGATGATGATGATGATGATGATGATGATG 448
412 CCTCATGATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAG 471
449 AAGAGTTGATCACTGCTTATGATCACTGATGATGATGATGATGATGATGATGATGATGAT 508
472 AAGAGTTGATCACTGCTTATGATCACTGATGATGATGATGATGATGATGATGATGATGAT 531
509 TAGAATCAAGCCTCAACCAAGCTTAATAGGATTAAGCACTTTGATGATGATGATGATGATGAT 568
532 TAGATGCTGCTTATTAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
569 ATTGATGAGCAGTCAAGCACTTAAGCACTTTGATGATGATGATGATGATGATGATGATGAT 628
592 CACTGATGAGCAGTCAAGCACTTAAGCACTTTGATGATGATGATGATGATGATGATGATGAT 651
629 AAGTTTGTGATCACTGCTTATGATCACTGATGATGATGATGATGATGATGATGATGATGAT 688
652 ATTCTCTGATCAAGCACTTATGATCACTGATGATGATGATGATGATGATGATGATGATGAT 711
689 AAGCTAGTATGATTAATTAATCACTGATGATGATGATGATGATGATGATGATGATGATGAT 748
712 AAGCCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
749 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 808
772 TGAGG---TAACACAGCTATCAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTAT 828
809 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 868
839 GAAACAAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888
869 AACAAAGTTTGTGAGGACTTTTACAAAGCTTAATGATGATGATGATGATGATGATGATGATGAT 928

```

```

889 AATTGACTAGAGTGG---ATGACCAATTAATCAAGGACCAACGCAAAATTACAGTTTG 945
929 TTGCTTTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 988
946 TTGCTTTAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
989 ATGCTGATGACCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1033
1006 CCTCTGAAATATCCTTTAGCACTACACAGCTGAAATGCAACCAAGCTGCAT 1065
1034 ---TCACAGATGTTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090
1066 CTTTACAGATGTTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125
1091 GTTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
1126 GCTTTCGTTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
1148 AAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
1186 AAC-----TAGGATTAATGTTGTTTACATTAATTAATTAATTAATTAATTAATTAATTAAT 1239
1208 AAGTGAAGTTTGCACCTTCACTAGTCAATATATATATATATATATATATATATATATATATAT 1267
1240 AATTTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1299
1268 TTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1302
1300 AACTATGCTAGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334

```

RESULT 6

US-09-601-198-55/c  
Sequence 55, Application US/09601198

Patent No. 6531583  
GENERAL INFORMATION:

APPLICANT: Caselli, Gail H.  
APPLICANT: Chen, Ellison Y.

APPLICANT: Glass, Jennifer S.  
APPLICANT: Glass, John I.

APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot

TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA

TITLE OF INVENTION: UREAPLASMIC

FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198

PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189

PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181

SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 55

LENGTH: 3057  
TYPE: DNA

ORGANISM: Ureaplasma urealyticum  
US-09-601-198-55

Query Match 5.0%; Score 65.6; DB 4; Length 3057;  
Best Local Similarity 46.2%; Pred. No. 3.7e-05;

Matches 218; Conservative 0; Mismatches 254; Indels 0; Gaps 0;

```

444 AGCTGAACGTTTAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 503
2350 ATGAGCTAAATTTAAATTAAGCAGATTTAGCTAAATGAATGATGATGATGATGATGATGAT 2291
504 TAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563
2290 TAATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2231
564 CCCAAATTTAGTGAAGATCAAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 623
2230 AATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2171

```

Qy 624 CTTGAGGTTTGTGATCACTGCTTAATATCAAAATTCGCAATATTTAGTGAATCTATA 683  
Db 2170 TCCTAATAGAGTATGAGAACATTAAGACGAATTTTCAAGAAATTAATCTAATACAT 2111  
Qy 684 CAATTAAGCTGATGTTTAATTAATCAAAACATGATCCACTAAATGGGGAGCGCTTTT 743  
Db 2110 TAACAAAAAAATTAATTAATGATATGAGATTAACATCATCAAGATGAGAACTAATGA 2051  
Qy 744 AGATTCTAATGATTAATCAAGTTAATCGAAATTTAATAATACGTTATCAACATTTAA 803  
Db 2050 ACTAATCTTAATCAAAACAAATTAATGATTAATTTGATTAATGATTAATCACTAATCTTA 1991  
Qy 804 TGAACAAAAGCTAATGCTGATGATTAATCTAATGTTTAAATTAATTAATGATTAATCA 863  
Db 1990 TAGAAATTAATTTCTTCAAAAAAGATTTTAATTTCAATTAATTAATCTTTGAAAGATTCCT 1931  
Qy 864 TAATGAACAAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCA 915  
Db 1930 AAATCAATATATATGCTGATTCATTTATATATATCATCATGACCACTGCA 1879

## RESULT 7

US-09-417-485D-5  
; Sequence 5, Application US/09417485D  
; Patent No. 6541202  
; GENERAL INFORMATION:  
; APPLICANT: Long, David M.  
; APPLICANT: Metz, Anneke M.  
; APPLICANT: Love, Ruchelle A.  
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes  
; FILE REFERENCE: 47714-5009-US  
; CURRENT APPLICATION NUMBER: US/09/417,485D  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 10640  
; TYPE: DNA  
; ORGANISM: Plasmodium falciparum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (834)..(7385)  
; OTHER INFORMATION: TERT gene  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1821)..(1837)  
; OTHER INFORMATION: m at position 1821 = a or c; w at position 1837 =  
; OTHER INFORMATION: a or t. Xaa (amino acid) at position 330 = Leu or  
; OTHER INFORMATION: Ile; Xaa at position 335 = Asp or Gly.  
US-09-417-485D-5

Query Match 4.2%; Score 55; DB 4; Length 10640;  
Best Local Similarity 42.7%; Pred. No. 0.0075;  
Matches 336; Conservative 0; Mismatches 450; Indels 1; Gaps 1;  
Qy 111 TTTCTTTCACACCAAAATCTCTAGTAATAAACCTTATTTATTTATTTTATGCTAT 170  
Db 3102 TTTGTCTTTACCATCAATTTTGTAGATATAAAAGTTAAAAAAAGAAAGAACCCAAA 3161  
Qy 171 CTTTAAGATATAATATATATCTTAATATCTATGATAGAAAGAAATCATCTTAAGAC 230  
Db 3162 ATGATTAATGATCAAAATATGTTCAATGTAAGAAAGAAAGGAGAAATTAATTAATTA 3221  
Qy 231 TATTAGTTTGTAGTACACATCTCTTCTTAGCATTTGGGATTTTCTACCTGTATCTAT 290  
Db 3222 AAATATATATTTATTAAGAAATGAAAAAGACCTAATAATGATTAATTAATTAATAA 3281  
Qy 231 TACTAAAGATGAGAACCCAAATATATGCGCAATCCCAATTAAGAGAGCGCGCAATGA 350  
Db 3282 TTTTCCAAAAAATGATCCCTTAATAAAAAAATTAATTAATTAATCAATCAACGTAT 3341  
Qy 351 GTTAACAGATCTAATCAATGCTAAGCGATGACATTAAGCTTCACTACCAAGACTATGCCA 410

Db 3342 AATATATATTTATTAATAAGATATGAAAAAATCAAAAACTAATTAATTAAT 3401  
Qy 411 GATTGAAGCTAGATTATCATCTGCTTATAGTGAAGCTGAACGTTAACAATACCTTA 470  
Db 3402 AAAAGTATGATTAATTTATCAAAATTAAG-GAAATTAACAAAAAGTTTGAACCATTA 3460  
Qy 471 TGCAACATTAGAACCACTAATAAATGCTAATAACTAATTTAGTAATCAGCCATCAACCAAG 530  
Db 3461 TATTAAAAATTTTACTATTAATTAATTAATTAATTAATTTTGTCTTAATAAAAAATGTAAT 3520  
Qy 531 TAATAGGATTAACGAAGCTTTGATTAATGAACACCAAAATTTAGTGAACATCAAGC 590  
Db 3521 TCATATGAGATGCAAGCAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 3580  
Qy 591 ACTAAAAACACTTAAGAACAGCTGCTACTAATCTGAAGCTTTCATCACTGCTTA 650  
Db 3581 TTTTATTAATTTTGTCTCAAGAAAAAAGAACATATTAATTAATTTAGTTCCATTTT 3640  
Qy 651 TAATCAATTTGCAATTAATTTAGTATGATCTATACATTAAGCTAGTTAATTAATTA 710  
Db 3641 TCAAAATGAAAGATTAATTTATGTTAATCAATTTAATTAATTAATTAATTAATTA 3760  
Qy 711 AACATGATTCACATAATGGGGAGCGTTTATGATTTCAATGATTAATCAAGTTAA 770  
Db 3701 TATATATTAAGCAAAACAGTGAATGTTAATAAATTAAGATTAAGATTAATTAATTA 3760  
Qy 771 TCGAATATTAATTAATCGTTATCACTATTAATGAACAAAGCTAATGCTATGCTAT 830  
Db 3761 AATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3820  
Qy 831 ATCTAATAGTTTATTAATAAAGATTCATAATTAATTAATTAATTAATTAATTAATTA 890  
Db 3821 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3880  
Qy 891 TACAAC 897  
Db 3881 TAACAC 3887

## RESULT 8

US-09-601-198-39/c  
; Sequence 39, Application US/09601198  
; Patent No. 6531583  
; GENERAL INFORMATION:  
; APPLICANT: Caesell, Gail H.  
; APPLICANT: Chen, Elison Y.  
; APPLICANT: Glass, Jennifer S.  
; APPLICANT: Glass, John I.  
; APPLICANT: Heiner, Cheryl R.  
; APPLICANT: Letkowitz, Elliot  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
; FILE REFERENCE: UAB-13452/22  
; CURRENT APPLICATION NUMBER: US/09/601,198  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/073,189  
; NUMBER OF SEQ ID NOS: 181  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 1887  
; TYPE: DNA  
; ORGANISM: Ureaplasma urealyticum  
US-09-601-198-39

Query Match 4.1%; Score 53.6; DB 4; Length 1887;  
Best Local Similarity 43.4%; Pred. No. 0.011;  
Matches 349; Conservative 0; Mismatches 449; Indels 6; Gaps 2;  
Qy 82 TAAATTAATTAATAAATTAATAAATGTTTCTTATCAACCAAAATCTAGTAATA 141  
Db 1284 TTAGTTAATTAATAACAAAGCTAATTTTAAATACTAATATGAAGATTTTAAATTAATT 1225







APPLICANT: Jung, G nther  
APPLICANT: Kellner, Roland  
TITLE OF INVENTION: Biosynthetic Process For The Preparation  
TITLE OF INVENTION: Of Chemical Compounds  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,625  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0980002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-392-625-16

Query Match 4.0%; Score 51.6; DB 2; Length 8700;  
Best Local Similarity 45.2%; Pred. No. 0.037;  
Matches 189; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
QY 512 AATGAGCATCAACCAAGCTAATAGCATTAACCACTTTGATATGAAACCAACTT 571  
DB 4195 ATTGGAATTTAATTAAGTGAAGAGAAATTTTATTAATTAATGCGAAGTTAT 4254  
QY 572 TAGTTGAAGCATTAACCACTTAACCACTTTGAAACCAAGTCTAATACTTGAAG 631  
DB 4255 ATCGTAGTAATGACATAAGAAATTAATAATTTATGCTAACTTACCAATCTTAAA 4314  
QY 632 GTTTGTCATCAAGCTGCTATATCAAAATTCGCAATTAATTTAGTGAATCTATACATAAG 691  
DB 4315 ATGACTATGAATTTTAAAAAAGAAATTTCCCAATCTTCAATTAATTTCTATTTAATAAAA 4374  
QY 692 CTAGTAGTTTAATACTAAACATAGATCCATAATGGGGGAGCGTTTAGATTCTA 751  
DB 4375 TTAGTAATTTAGAAATCTTAAAAAGACATCAAAAAAGCTTATATTAATCTTCA 4434  
QY 752 ATGAGATTACTACAGTTATCGGAATATTATATATAGCTTATCAACTTAATGACAAA 811  
DB 4435 GGATTAATGGCAGTTTATTAACATGCGTGTATATGATATATCGTATTAATCTGAAA 4494  
QY 812 AGACTAATGCTATGCTATATATAGTTTATTAATAAAAGTATCAAAATATATGAAAC 871  
DB 4495 AAGAAAAATTTGTTTATCTATTTTATTAATGAATTTACAAAACTAAAAAATTTGGGATG 4554  
QY 872 AAAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAACTAGAGTTTGT 929  
DB 4555 GTTGTGATTAATTAATTAACATTAATAAAATTTTGAATAATTAATCACTTTTGT 4612

RESULT 11

US-08-466-961A-16  
Sequence 16, Application US/08466961A  
Patent No. 5843709  
GENERAL INFORMATION:  
APPLICANT: Enlian, Karl-Dieter  
APPLICANT: G tz, Friedrich  
APPLICANT: Schnell, No. 5843709bert  
APPLICANT: Augustin, Johannes  
APPLICANT: Engelke, Gernar  
APPLICANT: Rosenstein, Ralf  
APPLICANT: Kaletta, Cortina  
APPLICANT: Klein, Cora  
APPLICANT: Wieland, Bernd  
APPLICANT: Kupke, Thomas  
APPLICANT: Jung, G nther  
APPLICANT: Kellner, Roland  
TITLE OF INVENTION: Biosynthetic Process for the Preparation of  
TITLE OF INVENTION: Chemical Compounds  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,961A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/392,625  
FILING DATE: 22-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
FILING DATE: 30-APR-1992  
APPLICATION NUMBER: US 07/784,234  
FILING DATE: 31-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0980004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8700 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-466-961A-16

Query Match 4.0%; Score 51.6; DB 2; Length 8700;  
Best Local Similarity 45.2%; Pred. No. 0.037;  
Matches 189; Conservative 0; Mismatches 229; Indels 0; Gaps 0;  
QY 512 AATGAGCATCAACCAAGCTAATAGCATTAACCACTTTGATATGAAACCAACTT 571  
DB 4195 ATTGGAATTTAATTAAGTGAAGAGAAATTTTATTAATTAATGCGAAGTTAT 4254  
QY 572 TAGTTGAAGCATTAACCACTTAACCACTTTGAAACCAAGTCTAATACTTGAAG 631  
DB 4255 ATCGTAGTAATGACATAAGAAATTTATTAATTTATGCTAACTTACCAATCTTAAA 4314  
QY 632 GTTTGTCATCAAGCTGCTATATCAAAATTCGCAATTAATTTAGTGAATCTATACATAAG 691

Db 4315 ATGACTATGAAATTTTAAAAAGAAATTCGAAATCTGATGATTTCTATTTAATAAA 4374  
Qy 692 CTGTAGTTTAAATTAACCTAAACCTGATCCTAATATGGGGAACGCTTTAGATTTGA 751  
Db 4375 TTAGTATTTTGAAGAAATCTTAAAAAGACCTACAAAAAGCTATATCTTACCTGCTTA 4434  
Qy 752 ATGAGATTACTACGATTAATCGGATATTAATATGATGATGATGATGATGATGATGAT 811  
Db 4435 GGTATATTTGGCAGTTTATTAACATGCGTTGATATGAAATTTTGGTATTTAATCTGAAA 4494  
Qy 812 AGACTAATGCTGATGATTAATGATGATTTTATTTAAAAAGTATTCAAAAATTAATGAAC 871  
Db 4495 AAGAAAAATTTGTTTATCTATTTTATGAAATTCAAAACTAAAAAATATGCGAGG 4554  
Qy 872 AAGTTTGTAGGAGCTTTTACAAACGCTAATGTTCAACCTTCAACTGATTTTGT 929  
Db 4555 GTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4612

## RESULT 12

US-09-801-861-3/c  
Sequence 3, Application US/09801861

Patent No. 6452154  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CLO01098  
CURRENT APPLICATION NUMBER: US/09/801,861  
CURRENT FILING DATE: 2001-03-09  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 53332  
TYPE: DNA  
ORGANISM: Human  
US-09-801-861-3

Query Match Best Local Similarity 4.0%; Score 51.6; DB 4; Length 53332;

Matches 149; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

Qy 24 ATATCTTCTGCTTAAAAAACAACAATCTTCTAACAATAATCTTAAATTAATAGCCGTTA 83  
Db 31619 AAT 31560  
Qy 84 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 143  
Db 31559 AAT 31500  
Qy 144 CGCTTATTTAT-TTTATTTTATGTCATCTTTAAGATTAATTAATTAATTAATTTCTA 202  
Db 31499 TAT 31440  
Qy 203 TGAATTAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 262  
Db 31439 TAT 31380  
Qy 263 GCATTGGATTTCTAGCTGATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 316  
Db 31379 AAT 31326

## RESULT 13

US-09-601-198-51/c  
Sequence 51, Application US/09601198

Patent No. 6531583  
GENERAL INFORMATION:  
APPLICANT: Caselli, Gail H.  
APPLICANT: Chen, Ellison Y.  
APPLICANT: Glass, Jennifer S.  
APPLICANT: Glass, John I.

APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
TITLE OF INVENTION: UREALYTICUM  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 51  
LENGTH: 1851  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-51

Query Match Best Local Similarity 3.9%; Score 50.8; DB 4; Length 1851;

Matches 166; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

Qy 533 ATAGGATTAACACGCTTTGATATGATGATGATGATGATGATGATGATGATGATGATGAT 592  
Db 794 AAAAAATGAAGAAATTTCTACAGCAATATATATGTTTGAACACCAATTAATTAAGTT 735  
Qy 593 TAAAGACCTTTGAACACGCTGCTACTTACCTTGAAGTTTGTCTCACTGCTTATA 652  
Db 734 TAAATCAATCAATTAATTTGATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTC 675  
Qy 653 ATCAATTTGCAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 712  
Db 674 ATATCTTATGATGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 615  
Qy 713 CACTAGATCACTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 772  
Db 614 TTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 555  
Qy 773 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 832  
Db 554 ACTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 495  
Qy 833 CTATAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 890  
Db 494 AAAAACTATGCAATTTCAACATTCAGAGAGCTTTAAAGTTTATGAAGAGATTT 437

## RESULT 14

US-09-601-198-56/c  
Sequence 56, Application US/09601198

Patent No. 6531583  
GENERAL INFORMATION:  
APPLICANT: Caselli, Gail H.  
APPLICANT: Chen, Ellison Y.  
APPLICANT: Glass, Jennifer S.  
APPLICANT: Glass, John I.  
APPLICANT: Heiner, Cheryl R.  
APPLICANT: Lefkowitz, Elliot  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
TITLE OF INVENTION: UREALYTICUM  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 56  
LENGTH: 14066  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-56

Query Match Best Local Similarity 3.9%; Score 50.6; DB 4; Length 14066;

Best Local Similarity 44.5%; Pred. No. 0.066;  
Matches 200; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

QY 461 ATACCTTAATGCAACATTAGAACAACTAATAAAGGCTTAAGATTTAGAACGCA 520  
DB 10561 ATACATCAATTTACATTTAATAAGCGCTGTAAGAACAGATTTTTCAGACCATG 10502  
QY 521 TCACCAAGCTAATACGATTAAGACCTTTGATATGAACACCCAAATTTAGTTGAAG 580  
DB 10501 CTAATTTATTTATGATTTAAGATTTGATGACAAAATTCGATTTATGAGATGTTA 10442  
QY 581 CATACAAAGCACTAAAACCACTTTAGAACAAAGCTGCTACTACCTTGAAGTTTCAT 640  
DB 10441 AAATTTATATCAAAATTTGATTAACATAAAAGCACTCAAAATTTGATATGGAAG 10382  
QY 641 CAACGCTTAATTAATTAATTTGCAATATTTAGTGATCTATACATTAAGCTAGTACT 700  
DB 10381 TTGTTGCAAAATTAATAATTAATTTCAATTTGTTGTTTAAAGAAAAACAACATATG 10322  
QY 701 TAATTAATAAACAATAATCCATTAATGCGGGAACCTTTTATGATTTCTATAGATTA 760  
DB 10321 TTATTAATAAAGCTTGAAGCTTTAATAAGAGTGAAGATTTGTTATGATGATTTG 10262  
QY 761 CTACATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 820  
DB 10261 ATTACTAGATCAAAATTAATTTTATGATCTAATAAACAACAACATAGGCTC 10202  
QY 821 CTGATGCTATCTATATGTTTATTAATAAAGATTCATTAATTAATTAATTAATTAAT 880  
DB 10201 TAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10142  
QY 881 TAGGACTTTTACAAACGCTAATGTTCAA 909  
DB 10141 TAGTGATGATTTTCAAGATTAATCAACAA 10113

## RESULT 15

US-09-601-198-181

Sequence 181, Application US/09601198

Patent No. 6531583

GENERAL INFORMATION:

APPLICANT: Casseil, Gail H.

APPLICANT: Chen, Ellison Y.

APPLICANT: Glaes, Jennifer S.

APPLICANT: Glaes, John I.

APPLICANT: Heiner, Cheryl R.

APPLICANT: Lefkowitz, Elliot

TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA

FILE REFERENCE: UAB-13452/22

CURRENT APPLICATION NUMBER: US/09/601,198

CURRENT FILING DATE: 2000-12-08

PRIOR APPLICATION NUMBER: 60/073,189

NUMBER OF SEQ ID NOS: 181

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 181

LENGTH: 2043

TYPE: DNA

ORGANISM: Ureaplasma urealyticum

US-09-601-198-181

Query Match 3.8%; Score 50; DB 4; Length 2043;  
Best Local Similarity 42.3%; Pred. No. 0.061;

Matches 348; Conservative 0; Mismatches 465; Indels 9; Gaps 1;

QY 442 GAAGCTGAACAGTTAATCACTTAATGCAATAGAACAACTAATAATGCTTAA 501  
DB 1126 GGATATGAATTTATTAATAATGATCATATGCAATTAACAGAAATTTAATAAAT 1185  
QY 502 ACTAATTTAGATCA-----GCCATCAACCAAGCTAATACGATTAACGACTTTT 552  
DB 1186 ACCATTTTAGAACACAAAGCTGTGCAACCAAGATGATCATGATTTTAAAGTTAAAT 1245

QY 553 GATATGACACCCCAATTTAGTTGAAGCATACAAAGCCTAATAAACCACTTTAGAACAA 612  
DB 1246 GATATGACAGAGAAATGAAATTTTGGACCAATCTGATTAATAATTAAGAAATCAAGAT 1305  
QY 613 CGTGCTACTAATCTTGAAGTTTGTCACTACCTGCTTATTAATCAAAATTCGATTAATTTA 672  
DB 1306 TCTACAGTCTGTTGAATTTAATGATTTCACTGAAAGTACTCAAGATGATTAACGAT 1365  
QY 673 GTGATCTATACAAATAAGCTAGATGTTTAATACTAAACATGATCCACTTAATGGG 732  
DB 1366 GAAATGTGATTAATTTAATTAATAAGCAATACTAGTTTAAACGATCTTAAACAA 1425  
QY 733 GAACGCTTTAGATTTCTAATGAGATTAATCACTAAGTATCGAATATTAATAACGTTA 792  
DB 1426 GAGATCATCTCAAGAAATATTTATTAATAAGATCTTTCATCAATTCGATTTA 1485  
QY 793 TCAACTTAATGAACAAAGACTAATGCTGATGATTAATCTAATTAATTTAATAAATA 852  
DB 1486 GATGAATTTAATTAATGATGATGAAGAAATTAATTAACAAACGAACTAATTCATAT 1545  
QY 853 GTGATTCAAATTAATGAACAAAGTTTGTAGGACTTTTACAAACGCTAATGTTCAACT 912  
DB 1546 ATGATGATGATTAATGAAGAAATTAATGATGATGATTAATTAATTAATTAATTAAT 1605  
QY 913 TCAACTACAGTTTGTGCTTTTATGCTGATGATTAACCCGCTCAATTAATTAATGATGA 972  
DB 1606 CGAATTCAGATCAAGAAATTAATGATCATGATTAATCAAAATTAATTAATTAATTAAT 1665  
QY 973 AGAAGACCGTTGGAATGATGAACCTTCAAGTGAATTCCTTGCAACACGAATAGT 1032  
DB 1666 ATTAATACCAAAATTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1725  
QY 1033 ATCAAGATGTTTGTGATTTAATGTTTATGCTGGAACAAACGAAATGATTTAGT 1092  
DB 1726 TATGAACCTGAAGAAATGTTAATTTGAATTTGAATTAATTAATTAATTAATTAATTAAT 1785  
QY 1093 TTGACAACTATGCTCATCACTGTTATTTAATTTCCCTTAATGTTGTTAAGCA 1152  
DB 1786 TCAATTAACCAAGATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1845  
QY 1153 GCTGATGCTAATAACGTTGATTAACAAATTAATAATTAATTAATTAATTAATTAATTAAT 1212  
DB 1846 AAGAAATATCTTTCTTAACACCAAAAGAAATTAATTAATTAATTAATTAATTAATTAAT 1905  
QY 1213 GAGTTGCCACTTCACTAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1254  
DB 1906 GAATTAACACCTGCTGATTTTGAATTAAGACACAGAAATTT 1947

Search completed: August 25, 2003, 05:45:37  
Job time : 109 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 03:22:22 ; Search time 1244 Seconds  
(without alignments)  
2360.211 Million cell updates/sec

Title: US-09-901-572a-1

Perfect score: 1306  
Sequence: 1 aaaaacatcagatgttcaat.....taaacgtttcaccagctt 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 1517243 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCRUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028	78.7	1152	14	US-10-131-591A-1
2	1024.2	78.4	1371	9	US-09-147-052-1
3	1024.2	78.4	3261	9	US-09-147-052-3
4	921.6	70.6	1082	14	US-10-131-591A-24
5	73.4	5.6	4985	14	US-10-094-240-10
6	68.6	5.3	8136	12	US-10-311-455-128
7	68.4	5.2	11836	12	US-10-240-453-114
8	68.4	5.2	11836	12	US-10-239-676-102
9	68.2	5.2	8392	12	US-10-311-455-1463
10	68	5.2	12405	14	US-10-240-453-43
11	68	5.2	12405	14	US-10-239-676-35
12	66.4	5.1	6669	12	US-10-204-708-6
13	66.4	5.1	6669	12	US-10-311-455-166
14	65.6	5.0	5163	12	US-10-311-455-1221
15	65.4	5.0	8771	12	US-10-311-455-1797
16	64.8	5.0	7351	12	US-10-311-455-1

C 17	64.8	5.0	18512	12	US-10-311-455-950	Sequence 950, App
C 18	64.2	4.9	17594	12	US-10-311-455-1939	Sequence 1939, Ap
C 19	63.2	4.8	16766	12	US-10-311-455-2130	Sequence 2130, Ap
C 20	63	4.8	9254	14	US-10-240-453-92	Sequence 92, App1
C 21	63	4.8	9254	14	US-10-239-676-86	Sequence 86, App1
C 22	62.6	4.8	5883	12	US-10-311-455-326	Sequence 326, App1
C 23	62.6	4.8	5883	12	US-10-240-485-26	Sequence 26, App1
C 24	62.2	4.8	5678	12	US-10-311-455-1111	Sequence 1111, Ap
C 25	62.2	4.8	6061	12	US-10-311-455-114	Sequence 114, App
C 26	62.2	4.8	9728	12	US-10-311-455-1876	Sequence 1876, App
C 27	62	4.7	5520	12	US-10-311-455-1492	Sequence 1492, Ap
C 28	61.4	4.7	7001	14	US-10-172-086-60	Sequence 60, App1
C 29	61.2	4.7	5152	12	US-10-204-708-73	Sequence 73, App1
C 30	61.2	4.7	5421	12	US-10-311-455-105	Sequence 105, App
C 31	61	4.7	6831	12	US-10-311-455-1460	Sequence 1460, App
C 32	60.8	4.7	15387	12	US-10-311-455-157	Sequence 157, App
C 33	60.6	4.6	5930	12	US-10-311-455-1368	Sequence 1368, Ap
C 34	60.4	4.6	5145	12	US-10-311-455-321	Sequence 321, App
C 35	60.4	4.6	5145	12	US-10-240-485-17	Sequence 17, App1
C 36	60.4	4.6	5641	12	US-10-311-455-1369	Sequence 1369, Ap
C 37	60.4	4.6	11996	12	US-10-240-485-45	Sequence 45, App1
C 38	60.2	4.6	7461	12	US-10-311-455-1758	Sequence 1758, Ap
C 39	60.2	4.6	9515	12	US-10-240-453-182	Sequence 182, App
C 40	60.2	4.6	9515	14	US-10-239-676-160	Sequence 160, App
C 41	60	4.6	6050	12	US-10-311-455-1984	Sequence 1984, Ap
C 42	60	4.6	9510	12	US-10-240-485-115	Sequence 115, App
C 43	60	4.6	9964	12	US-10-311-455-71	Sequence 71, App1
C 44	60	4.6	17967	12	US-10-311-455-988	Sequence 988, App
C 45	59.4	4.5	11049	12	US-10-204-708-23	Sequence 23, App1

ALIGNMENTS

RESULT 1  
US-10-131-591A-1  
; Sequence 1, Application US/10131591A  
; Publication No. US20030059799A1  
; GENERAL INFORMATION:  
; APPLICANT: Nippon Zeon Co., Ltd.,  
; TITLE OR INVENTION: Modified DNA molecules, Recombinants and uses thereof  
; FILE REFERENCE: J209  
; CURRENT APPLICATION NUMBER: US/10/131,591A  
; CURRENT FILING DATE: 2002-08-15  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1152  
; TYPE: DNA  
; ORGANISM: Mycoplasma gallisepticum  
; FEATURE:  
; OTHER INFORMATION: TTM-1 gene (after EcoRI)  
US-10-131-591A-1

Query Match	78.7%	Score 1028;	DB 14;	Length 1152;
Best Local Similarity	100.0%;	Pred. No. 3.8e-170;		
Matches 1028;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	279	CTGTATGCTTACTTAAAAAGATGCNAACCCCAATATATGCGCAAAACCCCAATTAGAGC	338	
DB	6	CTGTATGCTTACTTAAAAAGATGCNAACCCCAATATATGCGCAAAACCCCAATTAGAGC	65	
QY	339	AGCGCAATGAGTTAACAATCAATGCTTAAGCGATGACATTAGCTGACTACA	398	
DB	66	AGCGCAATGAGTTAACAATCAATGCTTAAGCGATGACATTAGCTGACTACA	125	
QY	399	AGACTATGCCAAGTTAGAGTATGCTTATGCTTAATGTAAGCTGAACAGTTAA	458	
DB	126	AGACTATGCCAAGTTAGAGTATGCTTATGCTTATGTAAGCTGAACAGTTAA	185	
QY	459	CAATTAACCTTATGCACTTAAAGCACTTAAATGGCTTAAATTTAGATGAC	518	
DB	186	CAATTAACCTTATGCACTTAAAGCACTTAAATGGCTTAAATTTAGATGAC	245	

QY	519	CATCAACCAAGCTAATACGGATTAACGCTTTGGTAATATGAAACACCCAAATTTACTTGA	578
Db	246	CATCAACCAAGCTAATACGGATTAACGCTTTGGTAATATGAAACACCCAAATTTACTTGA	305
QY	579	AGCATACAAAGCACTAAAAACCACTTTAACAACAAGCTGTACTAACCCTGAAGGTTGTC	638
Db	306	AGCATACAAAGCACTAAAAACCACTTTAAGAACAAGCTGTACTAACCCTGAAGGTTGTC	365
QY	639	ATCAACTGCTTATTAATCAAAATTCGCAATTAATTTAGTGAATCTATACATTAAGCTATGAG	698
Db	366	ATCAACTGCTTATTAATCAAAATTCGCAATTAATTTAGTGAATCTATACATTAAGCTATGAG	425
QY	639	TTTAATTAACATAAAACCTAGATCCACTAATATGGGGAAAGCGCTTTAGATTCTTAATAGAT	758
Db	426	TTTAATTAACATAAAACCTAGATCCACTAATATGGGGAAAGCGCTTTAGATTCTTAATAGAT	485
QY	759	TACTACAGTTAATCCGAATATTATTAATAACGTTATCACTATTAATGAACAAAGACTAA	818
Db	486	TACTACAGTTAATCCGAATATTATTAATAACGTTATCACTATTAATGAACAAAGACTAA	545
QY	819	TGCTATGCAATATCAATATAGTTTATTAATAAAAGGATTCAAATTAATGAACAAAGCTT	878
Db	546	TGCTATGCAATATCAATATAGTTTATTAATAAAAGGATTCAAATTAATGAACAAAGCTT	605
QY	879	TGTAAGGACCTTTTACAAACGCTAATGTTCAACCTTCAACCTACAGTTTGGTGCTTTAG	938
Db	606	TGTAGGACCTTTTACAAACGCTAATGTTCAACCTTCAACCTACAGTTTGGTGCTTTAG	665
QY	939	TGCTATGTAACACCCGTCATTAATTAATATGCAAGAAGAGACCGTTGGATGCTGATGA	998
Db	666	TGCTATGTAACACCCGTCATTAATTAATATGCAAGAAGAGACCGTTGGATGCTGATGA	725
QY	999	ACCTTCAGTGAATTCCTTGGCAACAAGATGTATCCAGATGTTTTCTTGATTTATAG	1055
Db	726	ACCTTCAGTGAATTCCTTGGCAACAAGATGTATCCAGATGTTTTCTTGATTTATAG	785
QY	1059	TTTAGCTGAAACAAACAGAGATCAAAATTTAGTTTAGCACTAATGTCATCAACTGG	1118
Db	786	TTTAGCTGAAACAAACAGAGATCAAAATTTAGTTTAGCACTAATGTCATCAACTGG	845
QY	1119	TTATTTAATATTCCTTATTAAGTTGGTTAAGCAGCTGATGCTAATTAACGTTGATTA	1178
Db	846	TTATTTAATATTCCTTATTAAGTTGGTTAAGCAGCTGATGCTAATTAACGTTGATTA	905
QY	1179	ATACAAATTAATATGAAATGTTCCAACAGTTGATGTTGGCACTTCAACATGAGCAAA	1233
Db	906	ATACAAATTAATATGAAATGTTCCAACAGTTGATGTTGGCACTTCAACATGAGCAAA	965
QY	1239	TAAATACACAGCTAATCCAACTCCAGAGTTGATGATTAAGTTGGCTAAATTCGTTTT	1298
Db	966	TAAATACACAGCTAATCCAACTCCAGAGTTGATGATTAAGTTGGCTAAATTCGTTTT	1025
QY	1299	ATCAGGTT 1306	
Db	1026	ATCAGGTT 1033	

RESULT 2  
 US-09-147-052-1  
 ; Sequence 1, Application US/09147052  
 ; Patent No. US20010014335A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SAITOH, Shuji  
 ; APPLICANT: TSUKAH, Yoshiharu  
 ; APPLICANT: YAMAGIDA, NO. US20010014335A1001  
 ; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR  
 ; TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE  
 ; FILE REFERENCE: 981167  
 ; CURRENT APPLICATION NUMBER: US/09/147, /052  
 ; CURRENT FILING DATE: 1998-04-05  
 ; PRIORITY APPLICATION NUMBER: JP 08-103548  
 ; PRIORITY FILING DATE: 1996-03-29

	/	PRIOR APPLICATION NUMBER:	PCT/JP97/01084	
	/	PRIOR FILING DATE:	1997-03-28	
	/	NUMBER OF SEQ ID NOS:	9	
	/	SOFTWARE:	PatentIn Ver. 2.1	
	/	SEQ ID NO 1		
	/	LENGTH:	1371	
	/	TYPE:	DNA	
	/	ORGANISM:	hybrid	
	/	US-09-147-052-1		
Query Match		78.4%	Score 1024.2;	DB 9; Length 1371;
Best Local Similarity		99.7%;	Pred. No. 1.8e-169;	
Matches 1026;	Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
OY		278	GCTGATGTCTATTTCCTAAAAAGATGGAAACCCCAATATATGCGCAAAACCAATTAGAAG	337
Db		188	GCTGATGTCTATTTCCTAAAAAGATGGAAACCCCAATATATGCGCAAAACCAATTAGAAG	247
OY		338	CAGCGCGATGGAAGTTAACAAGTCTAATCAATGCTAAAGCAGTAGCATTAAGCTTGACTAC	397
Db		248	CAGCGCGATGGAAGTTAACAAGTCTAATCAATGCTAAAGCAGTAGCATTAAGCTTGACTAC	307
OY		398	AAGACTATGCCAAGTTGAAAGCTAGTTTATCATCTGCTTAATAGTGAAGCTGAACAGTTA	457
Db		308	AAGACTATGCCAAGTTGAAAGCTAGTTTATCATCTGCTTAATAGTGAAGCTGAACAGTTA	367
OY		458	ACAATAACCTTAATGCAACATTAGAACCACTAAAATGGCTPAAAAGTAAATTAGATCAG	517
Db		368	ACAATAACCTTAATGCAACATTAGAACCACTAAAATGGCTPAAAAGTAAATTAGATCAG	427
OY		518	CCATCAACCAAGCTPATATGGATAAAGCAGCTTTGATATGAAACCCCAAATTTAGTTG	577
Db		428	CCATCAACCAAGCTPATATGGATAAAGCAGCTTTGATATGAAACCCCAAATTTAGTTG	487
OY		578	AAGCATCAAAAGCACTAAAAACACTTAGAACAACGCGCTACACACCCTGAAGGTTGT	637
Db		488	AAGCATCAAAAGCACTAAAAACACTTAGAACAACGCGCTACACACCCTGAAGGTTGT	547
OY		638	CATCAACTGCTTATATCAAAATTGGCAATAATTAGTGATCTATACAAATTAAGCTACTA	697
Db		548	CATCAACTGCTTATATCAAAATTGGCAATAATTAGTGATCTATACAAATTAAGCTACTA	607
OY		698	GTTTAATATACCTAAAACACTGATGCATTAATGGGGAAACGCTTTGATTTCTAATAGA	757
Db		608	GTTTAATATACCTAAAACACTGATGCATTAATGGGGAAACGCTTTGATTTCTAATAGA	667
OY		758	TTACTACAGTTAATCGAATATTAATATATCGTTATCAACTAATTAAGAACCAAGAATA	817
Db		668	TTACTACAGTTAATCGAATATTAATATATCGTTATCAACTAATTAAGAACCAAGAATA	727
OY		818	ATGCTGATGCAATTCTAATATGTTTTATTAAAAAGTGATTCAAAATPATGAACAAAGTT	877
Db		728	ATGCTGATGCAATTCTAATATGTTTTATTAAAAAGTGATTCAAAATPATGAACAAAGTT	787
OY		878	TTGAGAGGACCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGGCTTTTA	937
Db		788	TTGAGAGGACCTTTTACAAACGCTAATGTTCAACCTTCAAACTACAGTTTGGCTTTTA	847
OY		938	GTGCTGATGTAACACCCGTCATTAATTAATATGCAAGAGGACCGTTTGAATGATG	997
Db		848	GTGCTGATGTAACACCCGTCATTAATTAATATGCAAGAGGACCGTTTGAATGATG	907
OY		998	AACCTTCAATAGAAATCTTGCAAAACGAAATGATCAAGATGTTTCTTGATTTATA	1057
Db		908	AACCTTCAATAGAAATCTTGCAAAACGAAATGATCAAGATGTTTCTTGATTTATA	967
OY		1058	GTTTAGGTGGAACAAACACGAAGTACCAATTTAGTTTGAACATATGTCATCAACTG	1117
Db		968	GTTTAGGTGGAACAAACACGAAGTACCAATTTAGTTTGAACATATGTCATCAACTG	1027
OY		1118	GTTATTATTAATTTCCCTTAATAGTTGTTAAAGCAGCTGATCTAATACGTTGATTAAC	1177
Db		1028	GTTATTATTAATTTCCCTTAATAGTTGTTAAAGCAGCTGATCTAATACGTTGATTAAC	1087

Query 1178 AATACAAATTAATGAAATGTTCAACAGTTGTCACCTTCAACTAGTCGA 1237  
Db 1088 AATACAAATTAATGAAATGTTCAACAGTTGTCACCTTCAACTAGTCGA 1147  
Query 1238 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTCCAAATGCTTT 1297  
Db 1148 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTCCAAATGCTTT 1207  
Query 1298 TATCAGGTT 1306  
Db 1208 TATCAGGTT 1216

RESULT 3  
US-09-147-052-3  
Sequence 3, Application US/09147052  
Patent No. US20010014335A1  
GENERAL INFORMATION:  
APPLICANT: SAITOH, Shuji  
APPLICANT: TSUZAKI, Yoshinari  
APPLICANT: YANAGIDA, No. US20010014335A1oru  
TITLE OF INVENTION: NOVEL FIRED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,  
FILE REFERENCE: 981167  
CURRENT APPLICATION NUMBER: US/09/147, 052  
PRIOR FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: JP 08-103548  
PRIOR FILING DATE: 1996-03-29  
PRIOR APPLICATION NUMBER: PCT/JP97/01084  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 3261  
TYPE: DNA  
ORGANISM: hybrid  
US-09-147-052-3

Query Match 78.4%; Score 1024.2; DB 9; Length 3261;  
Best Local Similarity 99.7%; Pred. No. 2,5e-169;  
Matches 1026; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Query 278 GCTGATGCTATTAATGAAATGATGCAACCCAAATTAATGAGCCAACTTAGAAG 337  
Db 2078 GCTGATGCTATTAATGAAATGATGCAACCCAAATTAATGAGCCAACTTAGAAG 2137  
Query 338 CAGCGCGAATGAGTTAAAGATCTTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 397  
Db 2138 CAGCGCGAATGAGTTAAAGATCTTAATCAATGCTAAAGCGATGACATTAGCTTCACTAC 2197  
Query 398 AAGACTATGCGCAAGTTGAAGTATGATCTGCTTATAGTGAAGCTGAACAGTTA 457  
Db 2198 AAGACTATGCGCAAGTTGAAGTATGATCTGCTTATAGTGAAGCTGAACAGTTA 2257  
Query 458 ACAATTAACCTTAATGCAATTAAGCAACTTAATGAGTAAATGCTAAATTAATGAGTACAG 517  
Db 2258 ACAATTAACCTTAATGCAATTAAGCAACTTAATGAGTAAATGCTAAATTAATGAGTACAG 2317  
Query 518 CCATCAACCAAGCTTAATGCGATTAAGCACTTTGATTAATGAACACCCAAATTTAGATTG 577  
Db 2318 CCATCAACCAAGCTTAATGCGATTAAGCACTTTGATTAATGAACACCCAAATTTAGATTG 2377  
Query 578 AAGCAATTAAGGCTTAATGAAACCACTTTGAACCAAGCTGCTAATGCTTGAAGTTGT 637  
Db 2378 AAGCAATTAAGGCTTAATGAAACCACTTTGAACCAAGCTGCTAATGCTTGAAGTTGT 2437  
Query 638 CATCAACTGCTTATATCAATTCGCAATTAATTTAGTGAATCTATACATTAAGCTAGTA 697  
Db 2438 CATCAACTGCTTATATCAATTCGCAATTAATTTAGTGAATCTATACATTAAGCTAGTA 2497  
Query 698 GTTATATATCAATTAACCTAGATCCATTAATGCGGGAACGCTTTAGATTCTAATGAGA 757

Db 2498 GTTATATATCAATTAACCTAGATCCATTAATGCGGGAACGCTTTAGATTCTAATGAGA 2557  
Query 758 TTACTACAGTTAATCGGAATTAATTAATAGCTTAATCAACTTAATGAACAAAGACTA 817  
Db 2558 TTACTACAGTTAATCGGAATTAATTAATAGCTTAATCAACTTAATGAACAAAGACTA 2617  
Query 818 ATGCTGATGCAATTAATTAATGTTTATTAATTAATGAAATGCAATTAATGAACAAAGTT 877  
Db 2618 ATGCTGATGCAATTAATTAATGTTTATTAATTAATGAAATGCAATTAATGAACAAAGTT 2677  
Query 878 TTGATGAGACTTTTACAAACCGCTAATGTTCAACCTTCAACTACAGTTTGTGCTTTA 937  
Db 2678 TTGATGAGACTTTTACAAACCGCTAATGTTCAACCTTCAACTACAGTTTGTGCTTTA 2737  
Query 938 GTGCTGATGTAACCCGTCATTAATTAATTAATGCAAGAGCCGTTTGAATGATGATG 997  
Db 2738 GTGCTGATGTAACCCGTCATTAATTAATTAATGCAAGAGCCGTTTGAATGATGATG 2797  
Query 998 AACCTCAAGTGAATTTCTTGCAACAGATAGTATCAAGATGTTCTTGATTTATA 1057  
Db 2798 AACCTCAAGTGAATTTCTTGCAACAGATAGTATCAAGATGTTCTTGATTTATA 2857  
Query 1058 GTTATGCTGAAACAAACAGAGTACCAATTTAGTTTACCACTATGCTTCATCACTG 1117  
Db 2858 GTTATGCTGAAACAAACAGAGTACCAATTTAGTTTACCACTATGCTTCATCACTG 2917  
Query 1118 GTTATTAATTAATTCCTTAATTAATGATGTTAAAGAGCTGATTAATGATGATGATTA 1177  
Db 2918 GTTATTAATTAATTCCTTAATTAATGATGTTAAAGAGCTGATTAATGATGATGATTA 2977  
Query 1178 AATACAAATTAATTAATGAAATGTTCAACAGTTGATGTTCCACTTCAACTAGTCGA 1237  
Db 2978 AATACAAATTAATTAATGAAATGTTCAACAGTTGATGTTCCACTTCAACTAGTCGA 3037  
Query 1238 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTCCAAATGCTTT 1297  
Db 3038 ATATATCTACAGCTTATCCAACTCCAGAGTTGATGATTAAGTTCCAAATGCTTT 3097  
Query 1298 TATCAGGTT 1306  
Db 3098 TATCAGGTT 3106

RESULT 4  
US-10-131-591A-24  
Sequence 24, Application US/10131591A  
Publication No. US20030059799A1  
GENERAL INFORMATION:  
APPLICANT: Nippon Zeon Co., Ltd.,  
TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof  
FILE REFERENCE: 4209  
CURRENT APPLICATION NUMBER: US/10/131, 591A  
PRIOR FILING DATE: 2002-08-15  
NUMBER OF SEQ ID NOS: 79  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 24  
LENGTH: 1082  
TYPE: DNA  
ORGANISM: Mycoplasma gallisepticum  
FEATURE:  
OTHER INFORMATION: Modified TTM-1 portion (downstream of BglI) of  
OTHER INFORMATION: PNZ40K-S  
US-10-131-591A-24

Query Match 70.6%; Score 921.6; DB 14; Length 1082;  
Best Local Similarity 98.0%; Pred. No. 1.3e-151;  
Matches 933; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Query 355 ACAATTAATCAATGCTTAAGAGATGACATTTAGCTTCACTACAGACTATGCAAGATT 414  
Db 1 ACAATTAATCAATGCTTAAGAGATGACATTTAGCTTCACTACAGACTATGCAAGATT 60  
Query 415 GAAAGCTATTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATACTTAAATGCA 474

```

Db      61 GAAGTAGTTTCACTCTGTTATAGTAGAGCTGAAACGTTTAACTAATCCTGCAGGCA 120
Qy      475 ACATTGAAACAATAAAATGGCTAAACTAATTGAAATCAAGCCATCAACCACTAAT 534
Db      121 ACATTGAAACAATAAAATGGCTAAACTAATTGAAATCAAGCCATCAACCACTAAT 180
Qy      535 ACGGATTAACGACTTTGATATAGTAACGCCAAATTTGTTGAAGCAACAAACGACTA 594
Db      181 ACGGATTAACGACTTTGATATAGTAACGCCAAATTTGTTGAAGCAACAAACGACTA 240
Qy      595 AAAACACTTTAGAACACAGTGTACTAATCTGGAAGTTTGCATCACTGTTAAT 654
Db      241 AAAACACTTTAGAACACAGTGTACTAATCTGGAAGTTTGCATCACTGTTAAT 300
Qy      655 CAAATTCGCAATATTAGTGTATCAATCAATAAAGCTAGTATTAACTAAACA 714
Db      301 CAAATTCGCAATATTAGTGTATCAATCAATAAAGCTAGTATTAACTAAACA 360
Qy      715 CTAGATCCACTAAATGGGGGAAAGCTTTAGATTCCTAATGAGATTACTACAGTTAATGG 774
Db      361 CTAGATCCACTAAATGGGGGAAAGCTTTAGATTCCTAATGAGATTACTACAGTTAATAG 420
Qy      775 AATATTAAATATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATCATCT 834
Db      421 AATATCCAGATATACGTTATCAACTATTAATGAACAAAGACTAATGCTGATCATCT 480
Qy      835 AATAGTTTATTAATAAAAGTATTCAAAATAATGAACAAAGTTTGTGAGGACTTTTACA 894
Db      481 AATAGTTTATTAATAAAAGTATTCAAAATAATGAACAAAGTTTGTGAGGACTTTTACA 540
Qy      895 AACGCTAATGTCTCAACTCTCAACTAGTCTTGTGCTTTAGTGTGATGATGACACCC 954
Db      541 AACGCTAATGTCTCAACTCTCAACTAGTCTTGTGCTTTAGTGTGATGATGACACCC 600
Qy      955 GTCAATTATTAATATGCAAGAAAGAACCGTTTGAATGTGATGAACCTTCAAGTAAAT 1014
Db      601 GTCAATTATTAATATGCAAGAAAGAACCGTTTGAATGTGATGAACCTTCAAGTAAAT 660
Qy      1015 CTTCGAACACGAAATATGATCAAGATGTTTCTTGATTTATAGTTTACGTTGAACAAAC 1074
Db      661 CTTCGAACACGAAATATGATCAAGATGTTTCTTGATTTATAGTTTACGTTGAACAAAC 720
Qy      1075 ACGAGTACCAATTAGTTTACGAACTATGCTCATCACTGTTAATTATTTCCCT 1134
Db      721 ACGAGTACCAATTAGTTTACGAACTATGCTCATCACTGTTAATTATTTCCCT 780
Qy      1135 TATAAGTTGTTAAAGCAAGCTGATCTATAAAGTTGATTAACAATTAATAAT 1194
Db      781 TATAAGTTGTTAAAGCAAGCTGATCTATAAAGTTGATTAACAATTAATAAT 840
Qy      1195 GGAATGTTCACAAAGTTGATGCTCCACTCACTAGTCAAAATATACAGCTAAT 1254
Db      841 GGAATGTTCACAAAGTTGATGCTCCACTCACTAGTCAAAATATACAGCTAAT 900
Qy      1255 CCAACTCCAGCAGTGTATGATTAAGTTGTTAAATCGTTTATCAGTT 1306
Db      901 CCAACTCCAGCAGTGTATGATTAAGTTGTTAAATCGTTTATCAGTT 952

```

```

RESULT 5
US-10-094-240-10/c
; Sequence 10, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWISSEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N8289
; CURRENT APPLICATION NUMBER: US/10/094,240
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 10/056,405
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649

```

```

; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4985
; TYPE: DNA
; ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match      5.6%; Score 73.4; DB 14; Length 4985;
Best Local Similarity 43.3%; Pred. No. 0.002;
Matches 392; Conservative 0; Mismatches 511; Indels 2; Gaps 1;

Qy      110 TTTTCTTATCAACCAAAATCTCTAGTAATAAACGCTATTATTTTATTTAGTCA 169
Db      3163 TATGTGTATTAACAGTTTACTGAAATTTTACTATTATTAATTAATTAATTAATTA 3104
Qy      170 TCTTTTAAGATATTAATATCTTAATATCTATGAATTAAGAAAGATCATCTTAAGA 229
Db      3103 TTTTATATATGATTAATCTGAAATTAATTAAGAAATTAATTAATTAATTAATTA 3044
Qy      230 CTATTAGTTGTATGATCAACATCTTTCTTACAGTTGGATTTCTAGCTGATGCTA 289
Db      3043 TAATTAAGATTAACAAACAAATTAATTAAGAAACAAACAAATTAATTAAGAAAT 2984
Qy      290 TTAATAAAGATGACCAACCAATTAATGAGCAACCAATTAAGAAAGCAAGCGGAAATG 349
Db      2983 TAATTAATTAACATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2924
Qy      350 AGTTAAGATCTATCAATGCTTAAGGATGACATTAGCTTCAAGACTATGCA 409
Db      2923 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2864
Qy      410 AGATTGAAGCTAGTTTATCATCTGCTTATGATGAGCTGAACAGTTTACATTA 469
Db      2863 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2804
Qy      470 ATGCAACTTAGACATTAATAAATGGCTAAATTAATTTAGATTCAGCCATCAACCAAG 529
Db      2803 AGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2744
Qy      530 CTAAATAGGATTAACGACTTTTGTATATGAAACCAACCAATTTAGTTGAAGCATACAAG 589
Db      2743 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2686
Qy      590 CACTTAACCACTTTGAAACACGCTCTACTAATCTTGAAGTTTGTCAATCACTGCTT 649
Db      2685 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2626
Qy      650 AATATCAAAATTCGCAATTAATTTAGTATCTATCAATTAAGCTAGTATTAATACTA 709
Db      2625 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2566
Qy      710 AAACACTAGATCCATAATGGGGAAAGCTTTTGAATCTAATGAGATTACTACAGTTA 769
Db      2565 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2506
Qy      770 ATCGGAATATTAATTAATGCTATCACTATTAATGAACAAAGACTTAATGCTGATGAT 829
Db      2505 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2446
Qy      830 TATCTAATAGTTTATTAATAAAGATGATTCAAAATTAATGAACAAAGTTTGTAGGAGCTT 889
Db      2445 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2386
Qy      890 TTACAAACGCTAATGTTCAACTTCAACTACAGTTTGTGCTTTAGTGTGATGAT 949
Db      2385 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2326
Qy      950 CACCCGCAATTAATTAATTAATGAAGAACCGTTTGAATGTGATGAACCTTCAAGTA 1009
Db      2325 CTAAATGTGTACTTAATAAAGTTTATCTTTCTTAAATTAATAAATCAATCCTAAGGA 2266

```



QY	1010	GAATT	1014
Db	2265	AGCTT	2261

## RESULT 6

```

US-10-311-455-528/c
Sequence 528, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEBENDROCK, Christian
APPLICANT: BEHLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of a Specific Gene
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 528
LENGTH: 8136
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-528

```

Query Match	5.3%	Score 68.6	DB 12	Length 8136
Best Local Similarity	44.1%	Pred. No. 0.015		
Matches	375	Conservative	0	Mismatches 471; Indels 4; Gaps 2;
QY	2	AAACATCAGATTGTTAATCTGATATCTTTCCTTAAAAAAACACAAAAATCTCTAACAA	61	
Db	1591	AAATATCATTTTTATTATTACTATATACGAAATATATTAACCCGAAATTAACCTTAACAT	1532	
QY	62	ATCTCAAAATTAATAGCCGTTAAATTAATCTAAAAAAATTTAAAAAAATGGTTTCTTATCA	121	
Db	1531	TTAATTAATTAATAAAAAAATCAAAAAATATATTAATTAATTTTCTCAATTA	1472	
QY	122	ACCAAAATTCCTGTAGTAATAACGCTTATTATTTTATTTTACTCATCTTTAAGTA	181	
Db	1471	AAAAATTAATTAACAATATTAATTAATTAATTAATTAACACCCCTGTATTATTAATTAACCA	1412	
QY	182	TAAATATATCTTAATATATCTATGAAATAGAAGAATCATCTTAAAGCTTTAGTTGT	241	
Db	1411	TATCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1352	
QY	242	TAGGTACAAATCCCTTCTTGAATGGAATTTCTAGCTGATGCTATTACTAAAAAG	301	
Db	1351	TAAATCAAAAAAATTAACGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1292	
QY	302	ATGCAAAACCCAAATTAATGCGCCAAACCCAAATTAAGAAGCGCGGAATGAGTTAACAGATC	361	
Db	1291	ACAAATCAAAATTAACAAACTTAATCAATTAATTAATTAATTAATTAATTAATTAATTA	1232	
QY	362	TAATCAATGCTAAAGCGATGACATTAGCTGACTACAGACTATAGCCAAAGATTGAAGCTA	421	
Db	1231	TTTAAAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1172	
QY	422	GTTATATCTGCTTAATGAGTGAAGCTGAACAGTTAACATTAACCTTAATGCAATTAG	481	
Db	1171	CAAAACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1112	
QY	482	AACAATTAATAATGGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	541	
Db	1111	TAAATCTCAAAAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1053	

Qy	542	AAACGCTTTGATATATGAACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACA	601
Db	1052	AAAAATTAATTAATTAATTTTAACTAATTTAACCAACAAACAAAAAACTATTATTAATT	9939
Qy	602	CTTTAGACACAGTGCT---ACTAAGCTTGAAGTTGTGCATCAAGCTGTATATACAA	6586
Db	992	TTTCTCAAAAAATTAATTAACAAAAAAATTAACCTATTTAAAAATTTAAAAA	9338
Qy	659	TTTCGATTAATTAGTGATCTATACAAATTAAGTAGTTAATACCTAAACACTAG	718
Db	932	TTAATAATATTATATTAATATCCATATACATCTAATTTTAAAAAATCCTATATAAAACTA	8738
Qy	719	ATCCACTAAATGGGGAGCGCTTTAGA TTCTAAATGAGATTACTACGTTATCGGATA	778
Db	872	TTAACTTAACAATTACGACCTCTATATATACTTTTAAAAATCTTAATTAATAATCCATATAA	8138
Qy	779	TTAATAATACGTTATCACTATTATATGAACAAAAAGCTAATGCTGAATCTATATCTATA	838
Db	812	AAAATATTTTATATATTAATACTATTAATACATACTAAACATTAATAAACTATTAAACCAA	7538
Qy	839	GTTTATTTAA	848
Db	752	AAATTATTTAA	743

## RESULT 7

US-10-240-453-114/c  
Sequence 114, Application US/10240453  
Publication No. US20030148326A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
TITLE OF INVENTION: Transcription  
TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated  
TITLE OF INVENTION: with DNA Transcription  
FILE REFERENCE: 5013.1009  
CURRENT APPLICATION NUMBER: US/10/240,453  
CURRENT FILING DATE: 2002-10-02  
PRIOR APPLICATION NUMBER: PCT/EP01/03973  
PRIOR FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 350  
SEQ ID NO 114  
LENGTH: 11836  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (7603)  
US-10-240-453-114

	Query Match	5.2%	Score 68.4;	DB 12;	Length 11836;	
	Best Local Similarity	44.3%	Pred. No. 0.02;			
	Matches	372;	Conservative	0;	Mismatches 461;	Indels 6; Gaps 2
Qy	37	AAAAAAAAAACTCTTAACAAAATCCGTAATAAATTTTAAAGCGTTAATTAACTAAAA	96			
Db	4357	ATPAAAT	4299			
Qy	97	ATPAAAAAAATGGTTTTCTTATCAACAATAATTCCTAGTAATPAAACGCTTAATTAATTT	156			
Db	4297	ATPAAAT	4238			



US-10-311-455-1463/c  
; Sequence 1463, Application US/10311455  
; Publication No. US20030143606a1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIERENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting  
; FILE OF INVENTION: Cytosine methylation  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 1463  
; LENGTH: 8392  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-1463

Query Match 5.2%; Score 68.2; DB 12; Length 8392;

Best Local Similarity 45.3%; Pred. No. 0.019; Mismatches 443; Indels 16; Gaps 3;

Matches 380; Conservative 0; Mismatches 443; Indels 16; Gaps 3;

36 AAAAAAAAAAAGGTTTCTTACCAAAATCTTAATAATTAAGCCCTTAATTAATACTAAATAA 95  
6745 AAAAAAAAAAAGGTTTCTTACCAAAATCTTAAATTAATTAAGCCCTTAATTAATACTAAATAA 6686  
96 AATTAATAATAAGGTTTCTTACCAAAATCTTAAATTAATTAAGCCCTTAATTAATACTAAATAA 155  
6685 AAAAAAAAAAAGGTTTCTTACCAAAATCTTAAATTAATTAAGCCCTTAATTAATACTAAATAA 6626  
156 TTTATTTTATGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 215  
6625 AAAAAAAAAAAGGTTTCTTACCAAAATCTTAAATTAATTAAGCCCTTAATTAATACTAAATAA 6566  
216 AATCATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 275  
6565 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6506  
276 TAGCTGATGCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 335  
6505 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6446  
336 AGCAGCGGATGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 395  
6445 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6386  
396 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 455  
6385 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6330  
456 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 515  
6329 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6272  
516 AGCAGCGGATGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 575  
6271 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6222  
576 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 635  
6221 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6162  
636 GTATCAATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 695

Db 6161 ACTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6102  
Oy 696 TAGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 755  
Db 6101 AAAAAAAAAAAGGTTTCTTACCAAAATCTTAAATTAATTAAGCCCTTAATTAATACTAAATAA 6042  
Oy 756 GATTACTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 815  
Db 6041 AATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5982  
Oy 816 TAAATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 874  
Db 5981 TAAATCAATCTTAAGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5923

RESULT 10  
US-10-240-453-43/c  
; Sequence 43, Application US/10240453  
; Publication No. US20030148326A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIERENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
; TITLE OF INVENTION: Transcription  
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated  
; FILE OF INVENTION: With DNA Transcription  
; FILE REFERENCE: 5013.1009  
; CURRENT APPLICATION NUMBER: US/10/240,453  
; CURRENT FILING DATE: 2002-10-02  
; PRIOR APPLICATION NUMBER: PCT/EP01/03973  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 350  
; SEQ ID NO 43  
; LENGTH: 12405  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; NAME/KEY: unsure  
; LOCATION: (7895)  
US-10-240-453-43

Query Match 5.2%; Score 68; DB 12; Length 12405;  
Best Local Similarity 43.3%; Pred. No. 0.024;  
Matches 365; Conservative 0; Mismatches 475; Indels 2; Gaps 1;  
35 TAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 94  
9308 TAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9249  
95 AATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 154  
9248 AATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9189  
155 TTTATTTTATGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 212  
9188 TTTATTTTATGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9129  
213 AATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 272  
9128 AATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9069  
273 TTTATTTTATGATCTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 332

Db		9068 TCATTAATCTTGCAGCAATTAACAAATTATTAACGACAATAAATCCACTAATTAAACAT	9067
Oy		333 AGAACGAGCGCAATGGAGTTTAACGATTTAATCAAATGAAGCATTAAGCTTC	332
Dd		9008 AAAAAATTAACATATTAATAAAAAAAAAATTTAAAACCAACTTAATAAAAAAAAAATTTAAATCCAAA	8945
Oy		393 ACTACAGAAGCTATGCCAATGGAATGGAAGCTGTATATCTTGCTTTATAGGAAGCTGAAC	452
Dd		8948 AATAAAAAAACAAGCTACCACTTAATCAAAAAATCAAAAAATCAAAAAATTAATAATCTTTAAA	8885
Oy		453 AGTTAACATTAACCTTAATGCAACTTAGAACACTAAAAATGCGTAAACTAATTTAGA	512
Dd		8888 TCCAAAAAATAAAAAAAAAATTAATAAATTAACCAATCAAAATATTAATCTATCAAAACAAAT	8829
Oy		513 ATCAGGCATCAACCAAGCTTAATAGGATTAACGCACTTTGATTAAGAACCCAAATTT	572
Dd		8828 AATAAAAACTAATCAATAATTTTACCTTTAAAAAATCAATAATACTTAATCAAACTA	8765
Oy		573 AGTTGAAGCATACAAGCACTTAATAAACCACTTAGAACAGCTGCTAACTTTGAAG	632
Dd		8768 TTTTAACAAAAAATAAATTAATCTTAATAAATAATTTCAAAAAAATAAATA	8709
Oy		633 TTGTGCATCACTGCTTTAATCAAAATGCAATTTAGTGATCTATCAATAAGC	692
Dd		8708 AAACTAATAATTAACCACTTTAAAAAATTTCTACTACAAAAAATAAATAAATAAATC	8649
Oy		693 TAGTAGTTTAATTAACATAAACACTGATTCCTAATAATGGGGAAAGCTTTAGATTCTA	752
Dd		8648 AATATTTTTAAAAAAGCTAATAAATCTCAAAAAAATTTTTTTTTTCTTTTAAAAATTA	8589
Oy		753 TGAGATTCTACAGTTAATCGAATATTAATATATCGTTATCACTAATTAATGAACAA	812
Dd		8588 AACATATTTCAACATTAATAATTAATAAATAAATAATATCTAATTAATAAATAAATAATCAT	8529
Oy		813 GACTAATGCTGATGATTAATCTAATAGTTTATTAATAAATAAGATTCATAAATTAAGAA	872
Dd		8528 AACATATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	8469
Oy		873 AA 874	
Db		8468 TA 8467	

RESULT 11  
US-10-239-676-35/c  
Sequence 35, Application US/10239676  
Publication No. US20030082609A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
FILE REFERENCE: 5013.1003  
CURRENT APPLICATION NUMBER: US/10/239.676  
PRIORITY FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: PCT/EPO/013968  
DE 10019058.8  
DE 10019173.8  
DE 10032529.7  
DE 10043826.1  
PRIORITY FILING DATE: 2001-04-06  
2000-04-06  
2000-04-07  
2000-06-30  
2000-09-01  
NUMBER OF SEQ ID NOS: 228  
SEQ ID NO 35  
LENGTH: 12405  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

```

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (7895)
US-10-239-676-35

Query Match      5.2%; Score 68; DB 14; Length 12405;
Best Local Similarity 43.3%; Pred. No. 0.024;
Matches 365; Conservative 0; Mismatches 475; Indels 2; Gaps 1

QY 35 TAAAAAACACAAAATCTTCTACACAAAATCCATAATAATAGCCGTAAATTAACATAA 94
Db 9308 TTACACTTCTAATATATATTTTAAAAATATACAAACAACAAATCCCTAAAAAATACA 9244
QY 95 AAATTAAAAAAATGGTTTCTTATCAACAAAATCTCTAGTAATAACGCTTATTTAT 154
Db 9248 AAATTAACCTTCAATTTTATATAAAAACAATTCAATTAATATATACTACAAAAAACCA 9189
QY 155 TTTTATTTTATAGTCATCTTTTAAGTATATAATATATCT--TATATTTCTATGATAGAA 212
Db 9188 TTTTAAAAACATAAAAATATAAAAAACTTAAACCTTTAACTTCAACATCCACAAAAA 9129
QY 213 AAGATATCATCTTAAAGACTATATAGTTGTATGATGACATCTCTTTCTAGCATTTGGAT 272
Db 9128 CAAAACCACTAAACAGAAAATATATCTTAAAAAATAATCCAACTTAAAAATATTAAT 9068
QY 273 TTTAGCTGTATGTCTATTAACAAAAAGATGCAACCCAAATATATGCGCAACCAATT 332
Db 9068 TCAATATATCTTCAACGAAATAAACAATTTATACCAAAAATATAAATCCACTAATTTACAT 9009
QY 333 AGAAGACACGGGAATGAGAGTTAACAGATCTATATCAATGCTTAAAGCATGACATTAAGCTTC 392
Db 9008 AAAAAATTAACATATATATAAAAAAATTTAAAAACAACCTCTAAAAAAATTTAAAAATCCAAA 8949
QY 393 ACTACAAACTATGCGCAGATTGAAGTATGAGTTATTCATCTGCTATATATGAGGTGAAC 452
Db 8948 AATAAAAAAACACCTTACCACTAAATATCAAAAAAATCAAAAAATATATTAATCTTTAAA 8889
QY 453 AGTTAACATTAACCTTAAATGCAACATTAAGAACAACTAAAAATGGCTAAAACTAAATTAGA 512
Db 8888 TCCAAAAAATAAAAAATATAAAAAAATATTAACCATATCAATATTAATATCAATCAAAATTT 8829
QY 513 ATGCGCCATCAACCAAGCTATATAGGATTAATAAGACATTTTGATATATGAACCCAAATTT 572
Db 8828 AATAAAACTATCAATATATTTTACCTTTAAAAAATCATATATTAACCTTATATCAAAACTA 8769
QY 573 AGTTGAAGCATACAAAGACATTAATAAACCACTTTGAACAAGCTGCTCTATACCTTGAAG 632
Db 8768 TTTTATACAAAAAATTAATACTTAAAAATCTTATATAAAAAATTTCAAAAAAATAAAAAA 8709
QY 633 TTTGTATCAACGCTGTTATATCAAAATTCGCAATATTTAGTGATCTATACATTAAGC 692
Db 8708 AAAACTAAAAATAAACAACCTTTAAAAAATCTACTACAAAAAATAAAAAAATAAAAAAAC 8649
QY 693 TAGATAGTTTATTACTAAAAACATAGATCCACTAAATGGGGGAACGCTTTTATGATTTAA 752
Db 8648 AATATTTTAAAAAAACATATAAAATCTCCAAAAATTTTTTTTTTCTTTTAAAAAATAAA 8589
QY 753 TGAGATTAACACGATTAATGCGAATATTAATATAGCTTATCAACTTATATAGAACAAA 812
Db 8588 AACATTTTCAACATATATATATATCTATATAAAAAATATCTATATAAAAAATAAATATTT 8529
QY 813 GACTAATGCTGATGATCTATATATAGTTTATATAAAAAAGTATTCAAATATATGACA 872
Db 8528 AACATATCAAAAAAATAAAAAAATACTACTAAAAAATTAACAATTAATATATATATATATAT 8469
QY 873 AA 874
Db 8468 TA 8467

```



	Query Match	Similarity	Score	DB 12	Length	5163
Best Local	5.0%	45.5%	Pred. No. 0.047			
Matches	392	Conservative	0	Mismatches	459	Indels 11; Gaps 4;
Qy	28	CTTGCTTAAAAAACACAAATCTTTACCAAAATCCGAAATATATAGCGTTAAAT	87			
Db	4712	CTTAAACAATAAAAAATAAAAATATTATTAATTTATATTATTAATCTAATAAAAAT	4653			
Qy	88	AACTAAAAAATAAAAAATATGTTTTCTTATCAACCAAAATCTCTGATATAAAACGT	147			
Db	4652	AAATTCAAATTTTATPAAAAATATTATTAACCAATATTATPAAAAACATATAAATACCTAAC	4593			

```

RESULT 15
US-10-311-455-1797/C
; Sequence 1797, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424

```

```

; SEQ ID NO 1797
; LENGTH: 8771
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1797

```

```

Query Match      5.0%; Score 65.4; DB 12; Length 8771;
Best Local Similarity 48.2%; Pred. No.0.061;
Matches 284; Conservative 0; Mismatches 291; Indels 14; Gaps 3;

```

```

Qy 34 TTAACAAAAACAAATCTTCTAACAAATCTTAATATAATAGCGCTTAATTAACATA 93
Db 6791 TAAACAAAACTTAACATACTAACCAATAGATATACATACACATATATATCA 6732
Qy 94 AAAATTAACAAAAATGTTTTCTTATCAACCAAAATCTCTAGTATTAACGCTATTTA 153
Db 6731 ACACAAAAATATATATATATATATATATATATATATATATATATATATATAT 6672
Qy 154 TTTTATTTTATAGTATCTTTTAAAGATATATATATATATCT-----TATATCTATGA 205
Db 6671 TTCCAACTTATATCCAAACAAATAATATATATATATATATATATATATATATAT 6612
Qy 206 ATAGAAAAAGATCATCTTAAGACTATAGTTGTAGTACAAACATCTTTCTTAGCA 265
Db 6611 AAATTTTAACTATATTTTAAATAGCTAAATTAATTAATTAATTAATTAATTAAT 6552
Qy 266 TTGGATTTCTAGCTGATGCTATCTATTAATTAATTAATTAATTAATTAATTAAT 325
Db 6551 CACTATATTTCTTTTATATATATATATATATATATATATATATATATATATAT 6492
Qy 326 CCCAATTGAGACGCGGAGATGAGTAAAGATCTATCATGCTAAAGCATGACAT 385
Db 6491 TAAATTTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6432
Qy 386 TA--GCTTCACTACAGACTATGCGAAGATTGAAGCTAGTTATCATCTGCTTAGTGA 443
Db 6431 AAACCTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6372
Qy 444 AGCTGAAACAGTTAACATTAACCTTAATGCAACATTAGACACTAATAAATGGCTAAAC 503
Db 6371 AAATATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6316
Qy 504 TAAATTTAGATCAGCATCAACCAAGCTATATACGATATTAACGATTTGATTAATGAACA 563
Db 6315 TAACAAACAAACAGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6256
Qy 564 CCCAAATTTAGTGAAGCATACAAAGACATAAAGACACTTAAGACAA 612
Db 6255 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6207

```

Search completed: August 25, 2003, 06:06:33  
Job time : 1251 secs

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 03:07:11 ; Search time 3107 seconds  
(without alignments)  
10216.172 Million cell updates/sec

Title: US-09-901-572A-1  
Perfect score: 1306  
Sequence: 1 aaaaacatcagatgttcaat.....taaatcgtttatcaggtc 1306

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.4	6.1	1101	29	CNS0039G
2	79.2	6.1	1200	13	AL063921 Drosophila
3	78	6.0	1124	13	AL063921 Drosophila
4	77.2	5.9	1101	29	AL063921 Drosophila

Result No.	Score	Query Match	Length	ID	Description
5	77.2	5.9	1200	13	AL063921 Drosophila
6	77.2	5.9	1201	13	AL063921 Drosophila
7	76.4	5.8	1201	13	AL063921 Drosophila
8	76.2	5.8	1101	29	AL063921 Drosophila
9	75.8	5.8	1101	29	AL063921 Drosophila
10	75.6	5.8	1101	29	AL063921 Drosophila
11	75.4	5.8	1225	9	AL063921 Drosophila
12	75	5.7	1201	29	AL063921 Drosophila
13	74.8	5.7	1201	29	AL063921 Drosophila
14	74.2	5.7	998	13	AL063921 Drosophila
15	74.2	5.7	1201	13	AL063921 Drosophila
16	73.4	5.6	1146	29	AL063921 Drosophila
17	73	5.6	1201	13	AL063921 Drosophila
18	71.8	5.5	1201	13	AL063921 Drosophila
19	71.2	5.5	960	13	AL063921 Drosophila
20	71	5.4	781	29	AL063921 Drosophila
21	71	5.4	1056	13	AL063921 Drosophila
22	71	5.4	1091	13	AL063921 Drosophila
23	70.6	5.4	1056	13	AL063921 Drosophila
24	69.4	5.3	1092	29	AL063921 Drosophila
25	68.8	5.3	1101	29	AL063921 Drosophila
26	68.8	5.3	1074	29	AL063921 Drosophila
27	68.6	5.3	1201	13	AL063921 Drosophila
28	68.6	5.3	1201	13	AL063921 Drosophila
29	68.6	5.3	1187	9	AL063921 Drosophila
30	68.4	5.2	1098	13	AL063921 Drosophila
31	68.2	5.2	1200	13	AL063921 Drosophila
32	68.2	5.2	1178	29	AL063921 Drosophila
33	68	5.2	1178	29	AL063921 Drosophila
34	67.8	5.2	759	29	AL063921 Drosophila
35	67.8	5.2	1165	13	AL063921 Drosophila
36	67.8	5.2	1201	13	AL063921 Drosophila
37	67.8	5.2	1434	29	AL063921 Drosophila
38	67.6	5.2	576	29	AL063921 Drosophila
39	67.6	5.2	1131	29	AL063921 Drosophila
40	67.4	5.2	994	13	AL063921 Drosophila
41	67.4	5.2	1101	29	AL063921 Drosophila
42	67.4	5.2	1101	29	AL063921 Drosophila
43	67.2	5.1	1200	13	AL063921 Drosophila
44	66.8	5.1	839	13	AL063921 Drosophila
45	66.8	5.1	1101	29	AL063921 Drosophila

## ALIGNMENTS

RESULT 1  
CNS0039G  
LOCUS  
DEFINITION  
BACROBK10 of RPL1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION  
AL063921  
VERSION  
AL063921.1 GI:4941778  
SOURCE  
Drosophila melanogaster (fruit fly)  
ORGANISM  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pretergata; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>  
The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Ohsawa and Aaron Mammmer in Pieter de Jong's laboratory in the Department of

## COMMENT

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

**Source**

```
1.1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end = TET3"
```

BASE COUNT	201 a	64 c	131 g	202 t	503 others
ORIGIN					

ORIGIN

Query Match	6.1%;	Score 79.4;	DB 29;	Length 1101;
Best Local Similarity	19.1%;	Pred. No. 0.59;		
Matches 135;	Conservative 294;	Mismatches 272;	Indels 6;	Gaps 2

LOCUS BX437758/c 1200 bp mRNA EST 15-MAY-2003  
DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP0081B01  
ACCESSION BX437758  
VERSION 5-PRIME, mRNA sequence.  
KEYWORDS BX437758.1 GI:30773605  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1200)  
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue genoscope sequence ID : CS0CAP008CA010P1.  
Location/Qualifiers

**FEATURES**  
**Source**

## Sources

BASE COUNT	515 a	30 c	71 g	310 t	274 others
ORIGIN					

## ORIGIN

Query Match	6.1%;	Score 79.2;	DB 13;	Length 1200;
Best Local Similarity	30.8%;	Pred. No. 0.61;		
Matches 242;	Conservative 176;	Mismatches 343;	Indels 25;	Gaps 3;

[illegible]

```

RESULT 3
BX436282/c
LOCUS
DEFINITION
BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP001YC01
5-PRIME, mRNA sequence.
ACCESSION
BX436282
VERSION
BX436282.1 GI:30787521
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 1124)
Li, W.B., Gruber, C., Jessee, J., and Poyaves, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqr@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Peng Liang Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue, Genoscope, sequence ID : CS0CAP001AB01DPI.

```

```

Query Match          6.0%  Score 78;  DB 13;  Length 1124;
Beet Local Similarity 24.2%  Pred. No. 0.88;
Matches 203;  Conservative 255;  Mismatches 371;  Indels 10;  Gaps 2

Oy  83  AATGTAATAAAAAATTAATAAAATGTTTTCTTATCATCAACCAAAATCTCTAGTAATTA 142
      ||:::  ::||| |||||:::  ::::: || ||:::  ::  :: ||

```

RESULT 4	1101 bp	DNA	linear	GSS 03-JUN-1999
CNS003BD				
LOCUS				
DEFINITION				
	Drosophila melanogaster genome survey sequence TET3 end of BAC #			
	BAR08K08 of RFL1-58 library from Drosophila melanogaster (fruit			
	fly), genomic survey sequence.			
ACCESSION				
VERSION	AL064091.1	GI:4941847		
KEYWORDS				
SOURCE				
ORGANISM				
	Drosophila melanogaster (fruit fly)			
	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
	Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
	Genoscope - Centre National de Sequencage :			
	Submitted (02-JUN-1999)			

## COMMENT

BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))  
Determination of this BAC end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila  
melanogaster genome using these BACs. For further information  
please see <http://www.fruitfly.org> The BDGP Drosophila  
melanogaster BAC library was prepared by Kazuhiro Osoegawa and  
Aaron Mammeter in Pieter de Jong's laboratory in the Department of  
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
NY. The library is named RPCI-98 and was constructed by partial  
EcoRI digestion of Drosophila DNA provided by the BDGP from the  
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's  
P1 and EST libraries. A more detailed description of the library  
and how to order individual BAC clones, the entire library, or  
filters for hybridization from the BACPAC Resource Center can be  
found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

## FEATURES

## source

1. 1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR08K08"  
/clone\_lib="RPCI-98"  
/note="end : TET3"

BASE COUNT 395 a 120 c 103 g 334 t 149 others

## ORIGIN

Query Match 5.9%; Score 77.2; DB 29; Length 1101;  
Best Local Similarity 40.4%; Pred. No. 1.1;  
Matches 157; Conservative 47; Mismatches 185; Indels 0; Gaps 0;

QY 13 TTGTAATCTGATCTTCTGCTTAAACCAAACTCTTACCAAAATCTTAAATA 72  
DB 664 TTATTTAAWMAAATTAATAAAAAAAAAAAAAAAAAATATATAAAATTTATTAATW 723  
QY 73 ATACCGCTTAATTAACCTAAATAAATGATTTCTTATCACCATAATCT 132  
DB 724 TAAATATATATTTTAAATAATATTTTAAATAAATAATTTTATTTATTTAAATAAT 783  
QY 133 CTAGTAATAACGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 192  
DB 784 TATATTTAT 843  
QY 133 TATATTTCTATGATTAAGAAAGATCATCTTAAAGCATATGTTTGTAGTCAACA 252  
DB 844 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 903  
QY 253 TCCTTTCTAGCATTTGGGATTTCTAGCTATGTCATTAATAAATAAAGATCAACCA 312  
DB 904 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 963  
QY 313 AATAATGGCAAAACCAATTTAGACGCGCATGTGAGTTAACAATCTAATCTCT 372  
DB 964 AATATTTTAT 1023  
QY 373 AAAGCATGACCTTTAGCTTCACTACAAGA 401  
DB 1024 AATAACMAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1052

## RESULT 5

## LOCUS

BX437758 1200 bp mRNA linear EST 15-MAY-2003

DEFINITION BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP008YB01

5-PRIME, mRNA sequence.

## ACCESSION

BX437758

VERSION BX437758.1 GI:30773605

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1200)  
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Peng Liang Email : [liang@lifetech.com](mailto:liang@lifetech.com) URL :  
<http://fulllength.invitrogen.com> Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSOCAP008CA01QPI.  
Location/Qualifiers

## FEATURES

## source

1. 1200  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOCAP008YB01"  
/tissue\_type="THYMUS"  
/clone\_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

BASE COUNT 515 a 30 c 71 g 310 t 274 others

## ORIGIN

Query Match 5.9%; Score 77.2; DB 13; Length 1200;  
Best Local Similarity 30.6%; Pred. No. 1.1;  
Matches 256; Conservative 156; Mismatches 418; Indels 6; Gaps 1;

QY 16 TTATCTGATATCTTCTGCTTAAACCAAACTCTTCAACAAATCTTAAATA 75  
DB 365 TTATTTTAAWMAAATTAATAAAAAAAAAATGATTTCTTATCAACCAAAATCTCTA 424  
QY 76 AGCGTAAATTAATTAACCTAAATAAATGATTTCTTATCAACCAAAATCTCTA 135  
DB 425 TWTAAATAAAMWMDKKTITTTTCTTAAATAAATTTTCTTAAATAAATAATW 484  
QY 136 GTAAATAACGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 195  
DB 485 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 544  
QY 196 TATCTGATATTAAGAAAGATCATCTTAAAGCATATGTTTGTAGTCAACAATCC 255  
DB 545 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 598  
QY 256 TTCTTACATTTGGGATTTCTAGCTATGTCATTAATAAATAAAGATCAACCAAT 315  
DB 599 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 658  
QY 316 AATGCCCAAAACCAATTTAGAAAGCAGCGAAATGAGTTAACAATCTAATCTTAA 375  
DB 659 RDGARRTAAKATATATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAA 718  
QY 376 GCGATGACATAGCTTCACTACAAGACTATGCCAAGTTGAGTAGTTATCACTGCT 435  
DB 719 AGRAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 778  
QY 436 TATAGTGAAGCTGAACAGTTAACAATTAACCTTATATCAATTAAGAAACAATAATG 495  
DB 779 TTTTAAATAAAGTIDCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 838  
QY 496 GCTTAAATTAATTTAGATGACCATCAACCAAGCTAATACGATTAATAAACAATTTGAT 555  
DB 839 AAAAAAARAAATATATTTTATATATTAATAAATAAATAAATAAATAAATAAATAA 898  
QY 556 AATGAACCAACCAATTTAGTTGAAGCATCAACCAACCAATTTGAGCAACAGT 615  
DB 899 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 958  
QY 616 GCTACTAAGCTTGAAGTTGTCTCACTCAAGCTTATATCAAAATTCCTCAATTAATTAAGT 675

Db 221 AAAAAAAAAAAAAAAAAAATTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 280

Qy 327 CCAATTGAAGACAGCGGAAATGAGTTTAAAGATCTAATCAATCTAAAGCATGACATT 386

Db 281 AAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGMAAAACCCMAAAACCC 340

Qy 387 AGCTTCACTACAGACTATGTGCCAAGATTGAAGCTAGTTTATCATCTGTCTTATGTAAGC 446

Db 341 AACCAAAAAAAAAAAAAAAAAATATATAAAAAAAAAAAAAAAAAAAAAAAAAATWAAAAAAAAA 400

Qy 447 TGAACAGCTTACATATACCTTAATGCACTTGAACAATAAATGCTAAACCTAA 506

Db 401 AACCAATTTT 460

Qy 507 TTTGAATTCAGCCATCAACCAAGCTAAATACGATAAACGACTTTGATTAAGAACACC 566

Db 461 AATTA 520

Qy 567 AAATTGTGTAAGACTATCAAGACATTAACCACTTTAGAACACAGCTGCTAACT 626

Db 521 AAATTTTCCCAAAATTTTWTWA 580

Qy 627 TGAAGGTTTGCATCACTGCTTAATTAATCAAAATGCGAAATTAATTAAGATCTATCA 686

Db 581 WWWWTTTBTBAAAAADBRATTCYMAAACCCCTCTWAAAAAAACCYCAMHYMAAAAAA 640

Qy 687 TAAAGCTAGTATGTTAAATTAACCTAAAGACATGATCCATAATAGGGGAACGCTTTTGA 746

Db 641 AAAAAAAGCCCTTTTATTTTATTAAMAGCTTATGCGCHMMWAMCYTTTSBACCCCTWCT 700

Qy 747 TTTCAATGAGATTACTACAGTTAAATCGGAATATTATTAATTAATGCTTATCACTATTA 806

Db 701 TTTTCWCTCTWTTTCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 760

Qy 807 ACAAAAGCTATATGCTATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 866

Db 761 AA 820

Qy 867 TGAACAAAGTTTGTAGCG 885

Db 821 AAAAAAAAAARADWTTSTKG 839

RESULT 7

AL565455 1201 bp mRNA linear EST 12-MAY-2003

LOCUS AL565455 Homo sapiens PETA BRAIN Homo sapiens cDNA clone

DEFINITION CS0DF005Y018 3-PRIME, mRNA sequence.

ACCESSION AL565455

VERSION AL565455.2 GI:30549492

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

EUkaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li W.B., Gruber C., Jessee J. and Polayes D. Full-length cDNA libraries and normalization unpublished

On Feb 16, 2001 this sequence version replaced gi:12916848.

Contact: Genoscope

Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France

Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9232.f For more information about this cluster, see

http://www.genoscope.cns.fr/

cgf-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/

Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DF005BH09NP1.







/mol\_type="genomic DNA"  
/strain="REU927/4 GUTat 10.1"  
/db\_xref="taxon:5691"  
/clone="Sheared DNA-46123"  
/note="Vector: pUC18; Site 1: SmaI; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from a cloned population of  
Trypanosoma brucei (REU927/4 GUTat 10.1) was mechanically  
sheared to give a tight size distribution (approx 2 kb).  
The V + i method used for the library construction is  
described in detail in Smith, H.O. and Venter, J.C.  
(Making small insert libraries for whole genome shotgun  
sequencing projects. In Genome Sequencing: A Practical  
Approach, eds. M. Vaudin and B. Barrell, Oxford University  
Press, 1999)."

BASE COUNT 213 a 35 c 10 g 383 t

ORIGIN

Query Match 5.8%; Score 75.6; DB 28; Length 641;  
Best Local Similarity 46.0%; Pred. No. 2.2;  
Matches 255; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

QY 750 TAATGAGATTACTACAGTTATTCGGAATTTATTAATACGATCACTATTATGACGA 809  
DB 624 TAGTAACTACTACTAAATATATATATATATATATATATATATATATATATAT 565  
QY 810 AAGAGCTATGCGATGATTCATTCATAGTTTATTAATAAAGATTCATCAATATGA 869  
DB 564 TAAT 505  
QY 870 ACAAGCTTTTGAAGGACTTTTCAAGCTTAATGTTCACTTCAACTACAGTTTGT 929  
DB 504 TAAT 445  
QY 930 TCGTTTACGTCGATGTAACCCGTCATTTATTAATGCAAGAGACCGTTTGA 989  
DB 444 TAAT 385  
QY 990 TGGTGATGACCTTCAAGTGAATTCCTGCAACACAGAAATGATGACAGTGTCTTG 1049  
DB 384 TAAT 325  
QY 1050 GATTATAGTTTACGCGAACAACGAGTACCAATTTAGCTTTAGCACTATGCTCC 1109  
DB 324 TAAT 265  
QY 1110 ATCACTGCTTATTTATTTCCCTATATAGTTGTTAAAGCAGCTGATGCTATACGT 1169  
DB 264 TAAT 205  
QY 1170 TGAATTAACAATTAATAATATATATATATATATATATATATATATATATAT 1229  
DB 204 TAAT 145  
QY 1230 TAGTGCAATTAATCTACAGCTATCCACTCCAGAGTGTATGAGTTGCTAA 1289  
DB 144 TAAT 85  
QY 1290 AATGCTTTTATCAG 1303  
DB 84 AAGAGCTTTTGAAG 71

RESULT 11  
CC238324 1225 bp DNA linear GSS 12-MAY-2003  
LOCUS CH261-19212.RM1.1 CH261 Gallus gallus genomic clone CH261-19212,  
DEFINITION genomic survey sequence.  
ACCESSION CC238324  
VERSION CC238324.1 GI:30564987  
KEYWORDS GSS.  
SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
REFERENCE 1 (bases 1 to 1225)  
AUTHORS Kremitzki, C., Higgindotham, J., Wylie, K., Carter, J., McPherson, J.,  
Warren, W., Graves, T., Mardis, E. and Wilson, R.  
Gallus gallus BAC End Reads  
JOURNAL Unpublished  
COMMENT Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert length: 182000 Std Error: 0.00  
Seq primer: RM1 TACGACTCCTATAGGAGA  
Class: BAC ends  
High quality sequence start: 44  
High quality sequence stop: 100.  
Location/Qualifiers

FEATURES  
source 1..1225  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="CH261-19212"  
/sex="female"  
/cell\_line="UCD001, inbred 256"  
/note="Vector: pPARAC2.1, Site 1: EcoRI; Site 2: EcoRI;  
CH261 Female Chicken library - for library and clone  
ordering information: http://www.chori.org/bacpac"

BASE COUNT 548 a 32 c 23 g 379 t 243 others

ORIGIN

Query Match 5.8%; Score 75.4; DB 29; Length 1225;  
Best Local Similarity 38.8%; Pred. No. 1.8;  
Matches 330; Conservative 0; Mismatches 512; Indels 9; Gaps 2;

QY 4 AACATCAGATGTTAATCTGATATCTTGTCTTAAACCAAAATCTTCAACAAAT 63  
DB 259 AATTTTAAAAATAAAAATANNANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 318  
QY 64 CCTAATATATAGCCGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 123  
DB 319 AATAANNNAATNTTNTTAAATTAATTAATTAATTAATTAATTAATTAATTA 376  
QY 124 CAATATCTCTAGTAATTAAGCTTATTTATTTATTTATTTATTTATTTATTTAT 183  
DB 377 AATAANNNAATTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTAT 436  
QY 184 AATATATCTTATATTTCTATGATTAAGAAAGAAATCATCTTAAGACTATTAGTTG 243  
DB 437 ATTATNTTTTANNNTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 496  
QY 244 GGTACAACTCTCTTTAGCATTTGGATTTCTGCTGATGCTATGCTATTAATAA 303  
DB 497 ATAAAAAAATNTTTTNNNANNNNAATTAANNTNNNTAAATNTTANNTNANAAAA 556  
QY 304 GGAACCCCAATTAATGCGCAACCAATTAAGAGCGGCAATGAGCTTAACAGATCT 363  
DB 557 AAAATTAATTAATTAATTAATTTTAAANNTTATTAATTAATTAATTAATTAAT 616  
QY 364 ATCAATGCTTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 423  
DB 617 TAAATATTAATTAATTAATTTTNTTNTTNTTATTTTATTAATTAATTAATTA 676  
QY 424 TTATCATCTGCTTAATAGGAAGCTGAACAGTAACTAATCAATCAATCAATTA 483  
DB 677 NTATNTTAATTAATTAATTAATTTTATTTTNNAAAAATTAATTAATTAATTA 736  
QY 484 CAATTAATAATGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 543  
DB 737 NATATTAATAATAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 796



[illegible][illegible]



AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3370.r for  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE014CC03NP1&cluster=3370.r. Contact :  
Feng Liang Email: fliang@life-tech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE014CC03NP1.

FEATURES  
source

```
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE014RF05"
/tissue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-Oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
```

BASE COUNT 455 a 150 c 90 g 299 t 207 others  
ORIGIN

Query Match 5.7%; Score 74.2; DB 13; Length 1201;

Best Local Similarity 36.2%; Pred. No. 2.5;  
Matches 202; Conservative 80; Mismatches 275; Indels 1; Gaps 1;

```
1  AAAAAATCAGATCTTAACTGATATCTTGGCTTAATAAACAACAAATCTTCAACA 60
Db  640  AATATATATTTTAAATGTTGTTGATATTTTAAATTTTAAATTTTAAATTTTAA 699
Qy  61  AATCCTAAATTAATAGCGCTTAATTAATAAATAAATAAATAAATAAATAAATAA 120
Db  700  ATTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 759
Qy  121  AACCAAAATTCCTAGTATAAAGCTTATTTTATTTTATTTTATTTTATTTTATTT 180
Db  760  ATATATAATKTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 819
Qy  181  ATAAATATATCTTAATATCTATGATAAAGAAAGATCATCTTAAGACTATT-AGTT 239
Db  820  AATTTTATATATATTAATTTTAAATAAATAAATAAATAAATAAATAAATAAATA 879
Qy  240  GTTAGGTACACATCTTCTTCAAGTGGGATTTCTAGCTGATGCTATTTACTAAAA 299
Db  880  WTWWATATWTWTAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 939
Qy  300  AGATSCAAACCAATTAATGCGCAATTAAGAGCGGAGGAGGAGGAGGAGGAGGAG 359
Db  940  AATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 999
Qy  360  TCTAATCAATGCTAAGGAGTACATTAAGCTTCACTACAGACTATGCCAAGATTGA 419
Db  1000  AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1059
Qy  420  TAGTTATCATCTGCTTATAGTGAAGCTGAAACAGTTAACAATACTTAATGCAACT 479
Db  1060  TWWAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1119
Qy  480  AGAACAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 539
Db  1120  WAAANATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1179
Qy  540  TAAACGACTTTGATTA 557
```

Db 1180 WAAANATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1197

Search completed: August 25, 2003, 05:43:43  
Job time : 318 secs

**THIS PAGE BLANK (USPIC)**



Db 1 MHYPRNCFILVILVGTGNSPSTQNTVSREVSVOLSEESTFYLCPPVGVSTVRL 60  
 QY 61 E 61  
 Db 61 E 61

RESULT 2  
 YEEU\_ECO57 STANDARD; PRT: 2660 AA.  
 ID YEEU\_ECO57 Q8X2B9; Q8X2C0;  
 AC Q8X8V7; Q8X2B9; Q8X2C0;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yeeU.  
 GN 23135 OR EC52775/EC52776.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Escherichia.  
 CX NCBI\_TaxID=83334;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
 RL Nature 409:529-533 (2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohnishi E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Tida T., Takai H., Honda T., Sasaki A., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.,  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- SIMILARITY: Contains 16 Big-1 domains.  
 CC -1- SIMILARITY: BELONGS TO THE INTIMIN/INVASIN FAMILY.  
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
 frameshift in position 1315.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, AB005423; AAG57041.1; -  
 DR EMBL, AP002559; BAB36198.1; ALT\_FRAME.  
 DR EMBL, AP002559; BAB36199.1; ALT\_FRAME.  
 DR InterPro; IPR003344; Big-1.  
 DR InterPro; IPR003345; Intimin.  
 DR InterPro; IPR006001; PKD\_domain.  
 DR Pfam; PF02369; Big-1; 16.  
 DR PRINTS; PRO1369; INTIMIN.  
 DR SMART; SMO0634; BID\_1; 16.  
 DR SMART; SMO0089; PKD; 8.  
 KW Hypothetical protein; Repeat; Complete proteome.  
 FT DOMAIN 738 834 BIG-1 1.  
 FT 840 929 BIG-1 2.  
 FT DOMAIN 931 1033 BIG-1 3.  
 FT DOMAIN 1042 1132 BIG-1 4.  
 FT DOMAIN 1134 1236 BIG-1 5.  
 FT DOMAIN 1245 1335 BIG-1 6.

FT DOMAIN 1337 1439 BIG-1 7.  
 FT 1448 1539 BIG-1 8.  
 FT DOMAIN 1548 1632 BIG-1 9.  
 FT DOMAIN 1653 1750 BIG-1 10.  
 FT DOMAIN 1751 1855 BIG-1 11.  
 FT DOMAIN 1856 1957 BIG-1 12.  
 FT DOMAIN 1963 2056 BIG-1 13.  
 FT DOMAIN 2065 2156 BIG-1 14.  
 FT DOMAIN 2157 2252 BIG-1 15.  
 FT DOMAIN 2254 2355 BIG-1 16.  
 SQ SEQUENCE 2660 AA; 280062 MW; 01EB92A0BF5C09D2 CRC64;  
 Query Match 6.5%; Score 151; DB 1; Length 2660;  
 Best Local Similarity 23.0%; Pred. No. 1.8;  
 Matches 108; Conservative 70; Mismatches 188; Indels 104; Gaps 20;  
 QY 12 LIVILVGTGNSPSTQNTVSREVS---SVQLS---EEESTFYLCPPVGVSTVIRLEFGC 64.  
 Db 1131 IVNIAPDASNAQVTLNIPAOQVVTNNSDSVOLFTATVDPNNH---PVAGITV----- 1179  
 QY 65 MSTIKDANPNNGQTOLEAARMELTDLINKAMTLASLDQYAKIEASLSAYSAAETVNN 124  
 Db 1180 -----NFTMPQVVAANFTLEN--NGIATIQANGEAHVTLKGKAGTHVTATLGN 1227  
 QY 125 NLNATLEQLMAKATNLESAINQANT-----DKTT-----FDNEHPVLVAYK 166  
 Db 1228 N-NASDAQPTFPVADKDSAVVLOTSAELIIGNVDETTTLATVYKDFDN-----AVK 1279  
 QY 167 ALKTTLEQRAATLEGLSSTAVNQIRNLVDLYNKASSLIRKTYDPLNGTLL--DSNEIT 224  
 Db 1280 DLQVTF--STN---PADTQLSGKSNWTD-----SGVAETFK---GTVLGVHTAEAT 1324  
 QY 225 TVNRNINNTLSTNEQKTNDALNSFIKKVIONNEQSPFGTFNNAVQSPNSFVAFSA 284  
 Db 1325 LPNNNDNTKIVNIAPDASNAQVTLNIPAOQVVTNNSDSVOLATVXD--PSNHPVAGITV 1382  
 QY 285 DVFVNVKRYARRFTWNGDEPSSRIANTNSITDVSMTYISLAGNTKYQSPSNYGPSTGY 344  
 Db 1383 NFTMPQVVAANFTL---ENNGLAIQANGEAHVTLGKAGAGHTT-VTATLSNNTSDSQ 1437  
 QY 345 LYFPIKLVKADANNVGLQYKLNNGVQVYFPA-----STSANM 384  
 Db 1438 ---VFTEVADKTSALVVLQISKEITNGVDSALTITATVXDQEPNEVNNLPVTFSTASG 1494  
 QY 385 TTANPTPA-VDEIVAKIVISGLRFGQNTLESPTEGGMNKYAPMIGN 433  
 Db 1495 LTLPGESNTNESGIAQTILAGVAFGEQIVTASLANNGASDNKTVHETGD 1544

RESULT 3  
 NISP\_LACLA STANDARD; PRT: 682 AA.  
 ID NISP\_LACLA Q007596;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Nisin leader peptide processing serine protease nisp precursor  
 DE (EC 3.4.21.-).  
 GN NISP.  
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
 CX NCBI\_TaxID=1360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NITZO RS;  
 RX MEDLINE=93239683; PubMed=8478324;  
 RA van der Meer J.R., Polman J., Beerthuyzen M.M., Sizeren R.J.,  
 RA Kuipers O.P., de Vos W.M.;  
 RT "Characterization of the Lactococcus lactis nisin A operon genes  
 nisp, encoding a subtilisin-like serine protease involved in  
 precursor processing, and nisp, encoding a regulatory protein  
 involved in nisin biosynthesis.";  
 RT J. Bacteriol. 175:2578-2588(1993).

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=6F3;
RX MEDLINE=94213458; PubMed=8161176;
RA Engelle G., Gutowski-Eckel Z., Klesau P., Siegers K.,
RT "Regulation of nisin biosynthesis and immunity in Lactococcus lactis
6F3."
RL Appl. Environ. Microbiol. 60:814-825(1994).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=95357326; PubMed=7630881;
RA Stizen R.J., Rolfe H.S., Kuipers O.P., de Vos W.M.;
RT "Homology modelling of the lactococcus lactis leader peptidase Nisp
and its interaction with the precursor of the lantibiotic nisin."
RL Protein Eng. 8:117-125(1995).
CC -1- FUNCTION: CLEAVES THE LANTIBIOTIC NISIN PRECURSOR PEPTIDE.
CC -1- PATHWAY: Nisin biosynthesis; last step.
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
an amide bond (potential).
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L11061; AAA5200.1; -
DR EMBL; X76884; CAA54210.1; -
DR PIR; S44131; S44131.
DR HSSP; P29600; 1GCI.
DR MEROPS; S08.059; -.
DR InterPro; IPR006192; LPTXG.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; FALSE NEG.
DR PROSITE; PS50847; GRAM POS ANCHORING; FALSE NEG.
KW Hydrolyase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen;
KW signal.
FT SIGNAL 1 22
FT PROPEP 23 195
FT CHAIN 196 655
FT -----
FT PROPEP 656 682
FT ACT_SITE 259 259
FT ACT_SITE 306 306
FT ACT_SITE 512 512
FT SITE 652 656
FT MOD_RES 655 655
FT CONFLICT 500 500
SQ SEQUENCE 682 AA; 74767 MW; D5F29313F2983EC9 CRC64;
Query Match 6.5%; Score 150.5; DB 1; Length 682;
Best Local Similarity 21.0%; Pred. No. 0.34;
Matches 75; Conservative 60; Mismatches 131; Indels 91; Gaps 13;
QY 120 ETVNNNNATLEQLQMAKTNLESAINQANTDKTFEDNHPNVEAYKALKTTLEQRATNL 179
DB 38 ELINNSNAIISTEGSTDSINLGAQSPAYASTTRT---LDVTGAATLLQTSAAQK 93
QY 180 EGLSTAVNQIRNMLVDLYNKASSLITKTLPLNGTLLDSNEITVARNINTLSTINE 239
DB 94 EMKYSLSGTQVSE---FSKRDSTYTKENAPVSDLEEGSEVAVSSISQKN-KILDN 148
QY 240 QKTADALSNFPIKVIQNNNEOSFGVTTNANVPQSNYSFVAFSADVTVPVNYKARFTV- 298
DB 149 KKKRANFVTSPLIKKPSNSKDSAGVIDNS-----ASPLSYRKAKEVVS 193

```

```

QY 299 -----WNGDEPESRLANTNSITDVSWIYSLAGTNTKYQFSFNSGPS 341
DB 194 LRQPLKQKVEAQPLILSNSEKKAASYTNSHDVWDVQW-----DMKY---VTNNGES 243
QY 342 TGVLYPPKLVKAKADANNVGLQVTLNNGNVQVFPATSTSAN-----NT 385
DB 244 VA-LYOPSKKI-----SVGI---IDSGIMEHEDLNSLGNVPEKNLVPKCGFPNEBPDE 293
QY 386 TANPTPAVDEIKVAKIVLSGLRPGQNTIELSPGEGNMKNVAMIG-NILSSNEN 441
DB 294 TGNPSDIDV-----KMGHGTVEAQGITANGNIIIGVAGITVNIYRVEGEN 338
RESULT 4
AR56 CANAL STANDARD; PRT; 857 AA.
ID AR56 CANAL
AC P78586;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ARG5,6 protein, mitochondrial precursor [contains: N-acetyl-gamma-
glutamyl-phosphate reductase (EC 1.2.1.38) (N-acetyl-glutamate
semialdehyde dehydrogenase) (NAGSA dehydrogenase); Acetylglutamate
kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-glutamate 5-
phosphotransferase)].
GN ARG5,6
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
RX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 64385 / 1001;
RX MEDLINE=97195775; PubMed=9043106;
RA Negredo A., Monteoliva L., Gil C., Pla J., Nombela C.;
RT "Cloning, analysis and one-step disruption of the ARG5,6 gene of
Candida albicans."
RL Microbiology 143:297-302(1997).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
+ phosphate = N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate = ADP + N-acetyl-L-
glutamate 5-phosphate.
CC -1- PATHWAY: Arginine biosynthesis; second step.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
ACETYLGLUTAMATE KINASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAGSA
DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X98880; CAA67383.1; -
DR InterPro; IPR001048; Aa_kinase.
DR InterPro; IPR004662; AGLUK_kinase.
DR InterPro; IPR000706; AGPR_act_site.
DR InterPro; IPR006855; DUF619.
DR InterPro; IPR000534; Semialdh_dh.
DR Pfam; PF00696; aakinaase; 1.
DR Pfam; PF04768; DUF619; 1.
DR Pfam; PF01118; Semialdehyde_dh; 1.
DR Pfam; PF02774; Semialdehyde_dhc; 1.
DR Pfam; PF001765; AGPR_act_site; 1.
DR TIGRFAMs; TIGR00761; argB; 1.
DR PROSITE; PS01224; ARGK; 1.
KW Oxidoreductase; Transferase; Kinase; Arginine biosynthesis; NADP;

```

KW Mitochondrion; Multifunctional enzyme; Transf. peptide.  
 FT TRANSIT 1 2 MITOCHONDRION (POTENTIAL).  
 FT CHAIN ? 2 ACTIVITY:GLUTAMATE KINASE.  
 FT CHAIN ? 857 N-ACETYL-GAMMA-GLUTAMYL-PHOSPHATE  
 FT ACT SITE 669 669 REDUCTASE.  
 FT SEQUENCE 857 AA; 95066 MW; AEC2EAD8F8C4C71 CRC64;  
 SO  
 Query Match 6.4%; Score 149.5; DB 1; Length 857;  
 Best Local Similarity 22.4%; Pred. No. 0.51;  
 Matches 121; Conservative 72; Mismatches 189; Indels 157; Gaps 31;  
 QY 7 NCIFPLI-----VLYGTNSPTQNTVSREVSVQSESESTYLPCLPPVSGTVIR- 59  
 DB 104 SCLAPLVHGYPIYVHGT--GQINELLENGVEBEYIDGRIT-----NPKTMEVAK 156  
 QY 60 --LEFGCMSITKRD-----ANPNNGOTQLEAARMETLIDINAKAMTASLADYAK--1BA 110  
 DB 157 CLEQNLRLVTLAKETGVHARP-----ITAGVFAEYVDKDKYQVKGITSVNSPYEA 210  
 QY 111 SLSSAY-----SEAEVNN--NLAATL-----EOLAKAKTMBESAINQANTDK-- 151  
 DB 211 AINSGLPLITSLIAETSSQQLINADVAGELAREFEPKTIYVNEKGGIINGVTGEKV 270  
 QY 152 --TFEDNEHPLVEAKYKALKTLEQRATNLEGLSTAVYQINNLVDTLNKASSLITKTL 209  
 DB 271 SAINDEREVDL-----LKESWKYGTALK-----IKET-HDLIOHLPRESSVAIIDV 317  
 QY 210 DPLNGTLLDSNEITVTNN--IN-NLTSTINEOKTNADALSNEFIKKVIONNQSPFG 265  
 DB 318 NDLOKELEFDSGAGTLIRGRLINRNSLRDGC-----NPDILRNALLR----- 361  
 QY 266 TTTNNVQPSNVSFVAFSADVTPVNYKARRTVWNGDESSRIIANT----- 312  
 DB 362 --DEPIKTKVSVASYLKEFLDSVQPKS-----YGEDE-LEVLAIWEQNDKIPRLEDF 411  
 QY 313 -----NSIND-----VSWIYSLAGTNTKQFQSPSNVG--PSTGYLFPY-- 349  
 DB 412 LSSKGMANNVTDNFMNAIKDYSQLCWVNNENDANLPRYFSSKSGSPAKNGOILFWGL 471  
 QY 350 -----KLVKADANNVG--IQYKLNNG--NVQOVE-FATSTANNTTANPTPAVDK 397  
 DB 472 NIDASKLTIKEPSSSISGSSSKESGVTSAQCKRGFHSHTVRNT--NENPPLSEBK 529  
 QY 398 VA---KIYLSGLR--FGONTIELSVPTBGNKNKAPMIGNIYLSNENNAKIKGYAR 451  
 DB 530 QTERKKVALIGARGITGONLKLIT-----DNHPYLDISYSSRELEGGKLGQYNK 579  
 RESULT 5  
 CUT7 SCHPO  
 ID CUT7 SCHPO STANDARD; PRT; 1085 AA.  
 AC P24339;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Kinesin-like protein cut7.  
 GN CUT7 OR SPAC25G10.07C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91015362; PubMed=2145514;  
 RA Hagan I., Yamagida M.;  
 RT "Novel potential mitotic motor protein encoded by the fission yeast  
 RL Nature 347:563-566(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jørgensen K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Roben J., Grympe B.,  
 RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langier I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huzar S.M.,  
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM  
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;  
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO  
 CC FORM A SHORT SPINDLE THAT ELONGATES TO SPAN THE NUCLEUS AT  
 CC METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.  
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BINC  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X57513; CAA40738.1; -;  
 DR EMBL; Z70691; CAA94636.1; -;  
 DR PIR; T38378; T38378.  
 DR HSSP; P17119; 3KAR.  
 DR GeneDB SPombe; SPAC25G10.07C; -;  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin.1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.  
 KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;  
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.  
 FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).  
 FT DOMAIN 436 604 COILED COIL (POTENTIAL).  
 FT DOMAIN 715 740 COILED COIL (POTENTIAL).  
 FT NP BIND 897 955 COILED COIL (POTENTIAL).  
 FT REPEAT 987 998 ATP (BY SIMILARITY).  
 FT REPEAT 999 1010  
 FT MOD\_RES 1011 1011  
 FT CONFLICT 34 61 PHOSPHORYLATION (BY CDC2) (BY  
 FT FT SASNPRKRREPTIDTGYPRKSDTNSPT ->  
 FT FT LRAITGNDVSLTLTL (IN REF. 1).  
 SO SEQUENCE 1085 AA; 12213 MW; 566927787559D58 CRC64;  
 Query Match 6.3%; Score 147.5; DB 1; Length 1085;  
 Best Local Similarity 20.0%; Pred. No. 0.88;  
 Matches 109; Conservative 91; Mismatches 203; Indels 141; Gaps 25;



```

QY 17 YGINSSESTQNT-----SREVVSV-----QLSEESTYLCPPVGSYVIR 59
DB 554 YETWEAKITVAATLDSQYRESKEYIASLYEKLDRTERNNKNNENFM-----N 602
QY 60 LEFGCMSTTK---KDNPNNGQ--TOLEAAMELTDLINAKA-MTLASL---QDYAKIE 109
DB 603 LKFNILMTLRSFPHSFTDETGYFTLLNDFNASMBELINTSNOLLSMTKITEHPOLD 662
QY 110 ASLSAYSEATVNNNNATLEOLKMAKTLESAINOANTKTEDNEHPN-----VEA 164
DB 663 EALGSASSSCAVPSSLDLYSELKDSKSLDLAHSLODYSMSQGLGSGISELIEL 722
QY 165 YKALKT---LEGRATMLEGSSYAVNOIRNNLVLYNKASLLTKTLDPUNGTLDSN 221
DB 723 OKDKESYRQLVOELRSLYNLOHTHEESQKELMGVRNDIDALVYCTTSLDADILSD 782
QY 222 EIT-----TVRNNINNTLSTNEQ-KTNADL-----SNFICK--VI 256
DB 783 YISQKSFESKQODLIANIKIVSNFLQEQNESLYTQADILHSHLNDTNSIKRMEIM 842
QY 257 QNNBSFVGTFTNANVPQSNYSFVAFSADYTPVNYKARVTWNG---DEPSRILANT 312
DB 843 NNRSEEF---RNA-----ASQAEIVGANKERIQKTENGSQLDCKSAIHSNS 889
QY 313 NSITDVSNYSLA-----GTNTKYQ-----FSFSNYGPGSTGYLYPPYKLVKA 354
DB 890 RSMVD---HCLALAESQKQGVNLEVOITDLRLQKVKHSEEDNKEKHQOL---LDLLES 942
QY 355 ADANNVGL--QYKLANGVQVE---FATISANNTT-----ANFTPAVDELKV 398
DB 943 LVGNNDNLIDSIKTPHTELQKITDHVLKGTISLANHTMELLGLDESLCNETTIEDTSL 1002
QY 399 AKIVLSGLRFGQNTI-----ELSVPTGEGNMNVAPMIGNIYLSNENNADKIPG 448
DB 1003 VKLETTGTPTSKRELPAITPSTWTRDSSLIKETTNNLSDCKFVRETYTSSNQNEPDV-- 1060
QY 449 YRRP 452
DB 1061 YDKP 1064

RESULT 6
Y41_YEAST STANDARD; PRT; 719 AA.
ID Y41_YEAST STANDARD; PRT; 719 AA.
AC 003213;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 79.4 kDa protein in ALD2-DDR48 intergenic region.
GN YMR172W OR YMR8010.02.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churche C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagsels K., Iye G., Moulé S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XII."
RT Nature 387:90-93 (1997).
RL Nucleotide 387:90-93 (1997).
CC -1- SIMILARITY: LOW, TO YEAST MSN1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z49808; CAAB9905.1; -.
CC PIR: S55119; S55119.
CC TRASNPRC: TQ4601; -.
CC SGO: S0004783; HOTI.
CC DR GO: GO:0000228; C:nuclear chromosome, IDA.
CC DR GO: GO:0006972; P:hyperosmotic response, IGI.
CC K Hypothetical protein.
CC SEQUENCE 719 AA; 79415 MW; 4652BE93743D5A54 CRC64;

Query Match 6.1%; Score 142; DB 1; Length 719;
Best Local Similarity 21.2%; Pred. No. 1;
Matches 100; Conservative 69; Mismatches 175; Indels 128; Gaps 20;

QY 19 TNSSPSTQNTSREVVSVVOLSEESTFYLCPPVGSYVIRLERGCMSTTKDNPNNGQ 78
DB 150 TNSPS--NRISTDQLKIFQMDMS-----ARMTMESSEFNKSNKTAEQNTV 197
QY 79 TOLEAAMELTDLINAKAMTLA-----SLQDYAKIE-----ASLSAYSEATVNN 124
DB 198 LNLKQDYKVMANTLILKLVADQSPARPTNNQNKALIELNSISAVSAIYLOKQNNQ 257
QY 125 NLAATLEOLKMAKTLESAINQ-----ANTDKTFDNEHPNLEAVKALKTILEOR 175
DB 258 SGRQHTADLCTGDSNTHSGINQHRRTTGITIDVNTFQALNNQFSN-----ALNTILPDQ 311
QY 176 ATNLEGLSTAYNOIRNNLVLYNKASLLTKTLDPUNGTLDSNEITTVNRINNTL- 234
DB 312 QHN-----RNVVSQNTNQ--SLPNRLQGVIN-----TANQOQOVLI 348
QY 235 -STINQKTNADALS--NSFIKKYIQNNEQSFV---GFTTNANVPQSNYSFVA---FSA 284
DB 349 HNTTTHQVNRSPISFNASTDQPKLNPNGICRRRNTQSNNASTNDASAAQKPIA 408
QY 285 DVTVPNYKARVTWNGDEPS--SRILANTNSITDVSNYSLAGTNTKYQFSFSGYPPST 342
DB 409 -LSPLTNSHNTSTSMYNTSSIHSGVTSASNSFHDLN-----SLNPFQTTT 453
QY 343 GYLFPYKLVKADANNVGLQYKLANGVQVEPATISANNTTANFTPAVDELKVAKIV 402
DB 454 ALSPLSLALDNASPPPNQVPIPIINNTQQLPSFQILINDSTTS----- 498
QY 403 LSGLRFGQNTIELSVPTGEGNMNVAPMIGNIYLSNENNADKIPGRPGT 454
DB 499 -----EL-LPSGKSGVNT-----NIV---NRNPSLTPSPKPKMT 529

RESULT 7
SW1_YEAST STANDARD; PRT; 1314 AA.
ID SW1_YEAST STANDARD; PRT; 1314 AA.
AC P09547;
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription regulatory protein SW1 (SWI/SNF complex component SW1)
DE (transcription regulatory protein ADR6) (Regulatory protein GAA3).
GN ADR6 OR SW1 OR GAA3 OR YF016W OR LPA1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=97313271; PubMed=9169875;

```

RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Anseorge W.,  
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Bena V.,  
RA Boretan D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
RA Dietrich F.S., Deline H., DiPaolo T., Dubois E., Duesterhoeft A.,  
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,  
RA Huntic-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,  
RA Komp C., Kirdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
RA Marathe R., Messenguy F., Mewes H.-W., Mitterpat S., Moestl D.,  
RA Mueller-Auer S., Namath A., Nentwich U., Oetner P., Pearson D.,  
RA Petel F.X., Pohl T.M., Purnelle D., Schaefer M., Scharfe M.,  
RA Scherrens B., Schramm S., Schroeder M., Sdicu A.M., Tettelein H.,  
RA Utrerasazu L.A., Ushinsky S., Vierdeels F., Vissers S., Voss H.,  
RA Walsh S.V., Wambolt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
RA Zhong W.W., Zollner A., Vo D.H., Hant J.,  
RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XVI.",  
RL Nature 387:103-105(1997).  
[3]  
RP CHARACTERIZATION.  
RX MEDLINE=92154671; PubMed=1339306;  
RA Peterson C.L., Herskowitz I.,  
RT "Characterization of the yeast SWI1, SWI2, and SWI3 genes, which  
RL encode a global activator of transcription.",  
Cell 68:573-583(1992).  
CC -1- FUNCTION: INVOLVED IN TRANSCRIPTIONAL ACTIVATION. THE SWI/SNF  
CC COMPLEX IS REQUIRED FOR THE INDUCED EXPRESSION OF A LARGE NUMBER  
CC OF GENES. THIS COMPLEX ALTERS CHROMATIN STRUCTURE TO FACILITATE  
CC BINDING OF GENE-SPECIFIC DEDICATED TRANSCRIPTION FACTORS.  
CC -1- SUBUNIT: COMPONENT OF THE SWI/SNF GLOBAL TRANSCRIPTION ACTIVATOR  
CC COMPLEX.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: Contains 1 ARID domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial/  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X12493; CA31013.1; -  
CC EMBL; U33335; AAB68089.1; -  
CC PIR; S05728; TNEYR6.  
CC PDB; 1KKX; 04-DEC-02.  
CC PDB; 1KN5; 04-DEC-02.  
CC TRANSFAC; T01279; -  
CC SCD; S0005937; SWI1.  
CC InterPro; IPR001606; ARID.  
CC Pfam; PF01386; ARID.1.  
CC SMART; SM00501; BRIGHT.1.  
CC Transcription regulation; Activator; DNA-binding; Nuclear protein;  
CC Zinc-finger; 3D-structure.  
CC KW DOMAIN 5 65 ASN/THR-RICH.  
CC FT DOMAIN 337 385 GLN-RICH.  
CC FT ZN FING 1241 1258 C4-TYPE.  
CC SO SEQUENCE 1314 AA; 147938 MW; F442D5A82013CDBD CRC64;  
Query Match 6.1%; Score 141; DB 1; Length 1314;  
Best Local Similarity 19.9%; Pred. No. 2.5;  
Matches 79; Conservative 63; Mismatches 154; Indels 100; Gaps 17;  
QY 20 NSSPSTQVTSREVSVQLSEESTFYLCPPEVGYIRLEFGCMSTTKDAPNNGQT 79  
DB 13 NNTTTTNNNT 65  
QY 80 QLEAARMETLTLINAKANTLASLDQYAKIEASTLSAYSEATVYNNUNLATEQLKMATKN 139  
DB 66 GAGC---VDPFON-----FPQPKFPDQULDSNNNSNNNDNNNS--NTVASSTN 111  
QY 140 LES---AINQANTDKT-----FDNEHPNLVVEAYKALKTKTLEQRATNLEGLSTAVNQ 189

DB 112 FTSPFAYVNNNAAPNVGKAFNFIONQSPFNSPYDSNNNN-----TNINSLSPQAL-L 165  
QY 190 IRNVLVDLYNKASLLITKTLDPINGTLDSNEITTVNRNINNTLSTINEQKTNADALSN 249  
DB 166 AKNSIID-----SSNLPLOAQOQLYGG-----NNNNNSTGIANDVITPHFPTN 209  
QY 250 SFIKKVIONNQSVGVFTTANNVQ- SNYSFVAFSADVTPTNYKARRTVWNGGEPSSRI 308  
DB 210 --VQISQNSSSTSPN--TNSNSTPANNQQLPFPNNGASN-----NGNLTSQL 254  
QY 309 LANTNSITDVSWYISLAGTNTKYQSPFSNYPSTGYLYFPYKLVKADANNVGLQYKL-- 366  
DB 255 ISN-----YASNSMDRSSASANEFPVNT-----SDNNNNNNNNHNRN 292  
QY 367 -----NGNVOQVEFRTSANTTANPAPPAVDE 395  
DB 293 NSNNKTSNNNNVTVAVPATPATNTNSTNATVSE 328  
RESULT 8  
ID YS89 CAEEL STANDARD; PRT; 3178 AA.  
AC Q09624; Q09625; Q0969D4;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein ZK945.9 in chromosome II.  
GN ZK945.9/ZK945.10.  
OS *Caenorhabditis elegans*.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;  
OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
OX NCBI\_Taxid=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilkinson-Sproat J.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP REVISIONS.  
RA Dublin R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: Contains 1 GPS domain.  
CC -1- SIMILARITY: Contains 1 PLAT domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial/  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z48544; CAB70192.1; -  
CC EMBL; Z48582; CAB70192.1; JOINED.  
CC EMBL; Z48582; CAB70201.1; -  
CC EMBL; Z48544; CAB70201.1; JOINED.  
CC WormPep; ZK945.9; CE23697.  
CC InterPro; IPR002111; Cat\_channel\_TripL.  
CC InterPro; IPR005821; Ion\_trans.  
CC InterPro; IPR001024; Lipoxxygenase\_LH2.  
CC InterPro; IPR003915; PKD 2.  
CC InterPro; IPR000203; PKD\_Cys\_rich.  
CC Pfam; PF01825; GPS.1.  
CC Pfam; PF00520; Ion\_trans.1.  
CC Pfam; PF01477; PLAT.1.  
CC PRINTS; PR01433; POLYCYSTIN2.  
CC SMART; SM00303; GPS.1.  
CC SMART; SM00308; LH2.1.  
CC PROSITE; PS50095; PLAT.1.  
CC Hypothetical protein; Transmembrane.  
FT TRANSMEM 51 73 POTENTIAL.  
FT TRANSMEM 51 73 POTENTIAL.

```

MoJ. Microbiol. 15:39-54 (1995).
-1- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
-1- PM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
-1- SIMILARITY: TO YEAST SAG1.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ch/announce/ibb-ch/).
or send an email to license@ibb-eib.ch.
-----
DR EMBL, L25902; AAC41649.2; -.
KW Cell adhesion; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 1260 AGGLUTININ-LIKE PROTEIN 1.
FT DOMAIN 433 792 10 X 36 AA TANDEM REPEATS.
FT REPEAT 433 468 1-1.
FT REPEAT 469 504 1-2.
FT REPEAT 505 540 1-3.
FT REPEAT 541 576 1-4.
FT REPEAT 577 612 1-5.
FT REPEAT 613 648 1-6.
FT REPEAT 649 684 1-7.
FT REPEAT 685 720 1-8.
FT REPEAT 721 756 1-9.
FT REPEAT 757 792 1-10.
FT DOMAIN 983 1152 2 X 26 AA APPROXIMATE REPEATS.
FT REPEAT 983 1043 2-1.
FT REPEAT 1092 1152 2-2.
FT DOMAIN 399 404 POLY-THR.
FT DOMAIN 408 418 POLY-THR.
FT DOMAIN 450 455 POLY-THR.
FT DOMAIN 486 491 POLY-THR.
FT DOMAIN 522 527 POLY-THR.
FT DOMAIN 558 563 POLY-THR.
FT DOMAIN 594 599 POLY-THR.
FT DOMAIN 630 635 POLY-THR.
FT DOMAIN 666 671 POLY-THR.
FT DOMAIN 702 707 POLY-THR.
FT DOMAIN 738 743 POLY-THR.
FT DOMAIN 774 779 POLY-THR.
FT DOMAIN 874 877 POLY-SER.
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 687 687 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 723 723 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 820 820 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 886 886 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 918 918 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 973 973 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1068 1068 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1260 AA; 132641 MW; 763D1063A2354C24 CRC64;

Query Match 5.9%; Score 138; DB 1; Length 1260;
Best Local Similarity 21.2%; Pred. No. 3.5;
Matches 103; Conservative 75; Mismatches 189; Indels 118; Gaps 21;

18 GTNS-----SPSTQWTSREVVSVSLSEESTFLACPPVGSITVIRLEFGMSITKKDAN 73
Db 748 GTDIVIIRPPNPPIVITTEYWSQ--SFATTITVTAPPGGDIYIIE--SMSSSKISTS 802
QY 74 PN-----NGQTQ--LEAARMELTDLINAKAMTILASLDYAKIEASLSS 114
Db 803 SNDITSIIPFSRPHVYNSTTSDLPSTRESSMNTPTSISSDMLSS-----TT 851
QY 115 AYSGAELVNNNNLNTLEQLCKAKTNLESAINQANTDKTTPDNEHPNLVEAYKALKTILEQ 174
Db 852 LVYESETTTESICSDGKEC-----SRLSSESGIVTNPDSNBSISIVTSIVTPASTMSD 903

```

QY 175 RATNLGSLSTAVYQINRLNVLNPKAS-SLITKTLPLNGG-----TLDSNE 222  
 Db 904 SLSTGDSISATSSDNNVSKSVSTTERTSVTTIOTTEPNLSSVTSLSLTSSIPSVSESS 963  
 QY 223 ITTVNRINNTLSTINEOKTNAD-----ALNSFIKKVIONNQSQSVFGVTNANVOPSNY 277  
 Db 964 KVFSTNGDQSGSTHDSQSTSEIEIVTTSKVLPPVSSN-TDLTSEPTVTRREQPTL 1022  
 QY 278 SFV--AFSADVTVNKKVARTVWNGDEBPSRLIATNLSITDYSWISLAGINTKQVSEF 335  
 Db 1023 STTSNITIEDITL-----SPTGNDNDNTSS-----TNPVPTVA-TSTLASASEEDNKG 1071  
 QY 336 SNVPSGTGLYFFPKLVKADANNVGLQYKLNNGVQVEFATSTSNANTNP-----P 391  
 Db 1072 SHESASTS-----LKPSMGNSGL-----TTSTIEATTTSPTEAPSP 1109  
 QY 392 ANDEIVAKIVLSGLFPGQNTLELSPVBGNKKVAPMIGNIVYSSNNNADKIPGYRR 451  
 Db 1110 AVSSGTDVTEPTDREOPTLTS---TTSKTNSESAVAT-----QATNEN-----GGKS 1155  
 QY 452 PGTF 456  
 Db 1156 PSTDL 1160

RESULT 10  
 FNBA STAAU  
 ID FNBA STAAU STANDARD; PRT; 1018 AA.  
 AC P14738;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibronectin-binding protein precursor (FNBP).  
 FNBA.  
 OS Staphylococcus aureus.  
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OK NCBI\_Taxid=1280;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 8325-4;  
 RX MEDLINE=8909898; PubMed=2521391;  
 RA Signaes C., Raucet G., Joensson K., Lindgren P.-E.,  
 RA Anantharamiah G.M., Hoeseok M., Lindberg M.;  
 RT "Nucleotide sequence of the gene for a fibronectin-binding protein  
 RT from Staphylococcus aureus: use of this peptide sequence in the  
 RT synthesis of biologically active peptides".  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:699-703(1989).  
 CC -I- FUNCTION: THE ABILITY OF BACTERIA TO BIND FIBRONECTIN HAS BEEN  
 CC PROPOSED AS A VIRULENCE FACTOR ENABLING BACTERIA TO COLONIZE  
 CC WOUND TISSUES AND BLOOD CLOTS. BINDING OF PLASMA FIBRONECTIN TO  
 CC THE BACTERIAL SURFACE MIGHT BLOCK ADHESION RECEPTORS ON S. AUREUS,  
 CC THUS REPRESENTING AN IMPORTANT DEFENSE MECHANISM AGAINST TISSUE  
 CC INVASION.  
 CC -I- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by  
 CC an amide bond (Potential).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J04151; AAA26632.1; -  
 DR InterPro; IPR004237; Fn\_bind.  
 DR InterPro; IPR005877; Gpos\_Y5IRK.  
 DR InterPro; IPR001899; Gram\_pos\_anchor.  
 DR InterPro; IPR006192; LpXTG.  
 DR Pfam; PF02986; Fn\_bind; 1.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF04650; Y5IRK\_signal; 1.  
 DR TIGRfams; TIGR01167; LpXTG\_anchor; 1.

DR TIGRfams; TIGR01168; Y5IRK\_signal; 1.  
 DR PROSITE; P850847; GRAM\_POS\_ANCHORING; 1.  
 KW Cell wall; Peptidoglycan-anchor; Repeat; Signal.  
 FT SIGNAL 1 36  
 FT CHAIN 37 985  
 FT PROPEP 986 1018  
 FT REPEAT 545 574  
 FT REPEAT 575 604  
 FT DOMAIN 745 878  
 FT REPEAT 745 782  
 FT REPEAT 783 820  
 FT REPEAT 821 859  
 FT REPEAT 860 878  
 FT DOMAIN 879 948  
 FT REPEAT 879 892  
 FT REPEAT 893 906  
 FT REPEAT 907 920  
 FT REPEAT 921 934  
 FT REPEAT 935 948  
 FT SITE 982 986  
 FT MOD\_RES 985 985  
 SQ SEQUENCE 1018 AA; 111780 MW; 58175E0020E81F1F CRC64;  
 Query Match 5.9%; Score 137.5; DB 1; Length 1018;  
 Best Local Similarity 20.1%; Pred. No. 2.8;  
 Matches 88; Conservative 66; Mismatches 173; Indels 111; Gaps 19;

QY 19 TNSSPSTQNTVTSRREVSQVLSSEESTFYLCPPVGSVTIRLEFGCMSITKDNP---N 75  
 Db 81 TEGSNMNTQVTTTEAPVAPVQAPQTAQ-----PANIETKKE-----VKKEAKPVYKE 128  
 QY 76 NGQTQLEARMKELTDLINAKAMTLASLDYAKIFASISAYSEAEVYNNINLNTLEQKM 135  
 Db 129 TTQEQDMSGDQRQVYDLPKKATQNVAVETQVEVAPRTASBSKRVYRSADVAKASN 188  
 QY 136 AKTNLESAINQANTDKTTFD-----NEHPNLYEAYKALKTLEQRTNLEGSTAY-- 187  
 Db 189 AKVETGTDV-----TSKVTVEIGSLGHNNTNKVEPHAGQRAVLKYKLFENGHGGDFD 244  
 QY 188 ---NOIRNLVDLYNKA-----SLITKTLPLNGTTL-----DSNEIT----- 224  
 Db 245 FTLSNNVNTGVSTARKVPEIKNGSVVMATGEVLEGKIKYTFNDIEDKVDYTALEIN 304  
 QY 225 -----TVNRINNTL-STINEOKTN-----ADALNSFIKKVIONNQSQSVFGVTNNA 270  
 Db 305 LFLDPKTVQNTGQTLTSTINERQTSKELDVKYDGIQNYA-----NINGSIETFNKA 358  
 QY 271 NVQPSNYSFVAFSADVTPVNVYKTAARV-----WNGDEPSRLIATNLSITDVS-W 320  
 Db 359 N--NRSVAF--IKPNNGKTTSTVTTGLMKGNSQNGQPKVRLFEYLGNNEDIKAS 412  
 QY 321 IYSLAGINTKYQSFNSVGPSTGYLYFPYKLVKADANNVGLQYKLNNG-----NVQOVER 376  
 Db 413 VYANTDTTSKFEVTSNM-----SGNLNLQ-----NNGSYSLNIENLDK 451  
 QY 377 ATTSANNTANPTPAVD 394  
 Db 452 TYVHYDGEYLNIGTDEVD 469

RESULT 11  
 YKK1 CAEEL  
 ID YKK1 CAEEL STANDARD; PRT; 1010 AA.  
 AC P34278;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C02F5.1 in chromosome III.  
 OS Ctenorhadditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Ctenorhadditis.

```

XN [1] NCBI_TaxID=6239;
RN RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berke M.,
RA Bonfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Kershaw J., Kirsten J., Lalister N.,
RA Johnston L., Jones M., Keshav J., Kristen J., Lalister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Southammet E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston K., Watson A., Weinstock L., Wilkinson-Sprat C.,
RA Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL elegans."
RT Nature 368:32-38(1994).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC EMBL; L14745; AAA27916.2; .
DR WormPep; C02F5.1; C02450.
DR Hypothetical protein.
KW SEQUENCE. 1010 AA; 113231 MW; 15E19DD975624D94 CRC64;
SO -----

Query Match 5.9%; Score 137; DB 1; Length 1010;
Best Local Similarity 21.1%; Pred. No. 3;
Matches 107; Conservative 100; Mismatches 164; Indels 136; Gaps 28;

QY 21 SSPNQNTSEVSVVGLSEESFFYLCPPVGSVTYRLRGCGMSITRKD--ANNRN 76
DB NSPKRVASSNDLEKTIIE-ADKSKTKY--PSIADEVEDLDMD-NDITBQPCBAGNOON 417
QY 77 GQTOLEAA-RHELTDLINAKAM--TLASLDQYAKIEASISAYSEAE----- 120
DB DGLQIQKRDMDISVIRDSPAVDIMVAIFQSPARVYKIGANNISIDISQSIYVGDEMSIDE 477
QY 121 -----TYNNINATLEQLKRAKTNLESALINQANT--DKTFPNEHPN 160
DB TQNDGTLLPKRSNVEVITTDVYTSLEKQEEASNSVSMINSSVSHSEIDKKSF-----M 532
QY 161 LVEAVKA-LKTTLEGRATNLE--GLSSPAVYQIRNNVLDVYKASLITK----- 207
DB LIIEERAAWHSMSMDVDAQLEDDGSKTP-----VILASQASLAKPEPSALHNSGA 584
QY 208 TLD---PLNGDTLDSNEITTYNRRINNTLSTINEQ--KTNAADLNSFTKRVIONNEQ 261
DB TLNNSMELDNNTLTKMQITTC-----EDISWVHSIAVELNNSNKEQEGFGFETIQKNDT 639
QY 262 SFVG---TF-----TNANVPSTSYFVA-ESADYTPVNYKYA--RRTVWNGDEPS 305
DB SNTGANTFPQCHNETSQIMNNVDSAAVNTSKISTYSAFNLISINOSISKRRLSINARES 699
QY 306 SRIILANTSGITDVSMIYSLAGTN--TKYQPSFNSNGPSTGYLFKYLVKAAADANNVG 361
DB PRRAVLENS-----INSNGQTMALTEYR--QNKTMQTSIQDSMP-----SMLNSG 745
QY 362 LQYKLNNGVQVEFATSTSNNTTANPTPAVDIEIKVAKIYVLSGLRFGQNTIELSVPTGE 421
DB RDILAMNTSVR-----SPLNNSKTAAPGPBSL-----MSQN-VQLPPSPQ 786
QY 422 GNAKVAAPMIGN-ITLSSNENNAADKIP 447
DB FEMDPDFPAVVNVVYLTSEDPESTQHP 813

```

ID	Y109_YEAST	STANDARD	PRT:	995 AA.
AC	P40442;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Hypothetical 99.7 kDa protein in SDL1 5' region precursor.			
GN	Y11696 OR Y19402.07C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288c / AB972;			
RX	PubMed=9169870;			
RA	Churrier C.M., Bowman S., Badcock K., Bankier A., Brown D.,			
RA	Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,			
RA	Harris D.E., Horsnell T., Hunt S., Jagsels K., Jones M., Lye G.,			
RA	Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,			
RA	Skellton J., Smith V., Walsh S., Whitehead S., Barrett B.G.,			
RT	"write nucleotide sequence of Saccharomyces cerevisiae chromosome IX."			
RL	Nature 387:84-87(1997).			
CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.1ib-sib.ch/announce/">http://www.1ib-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@1ib-sib.ch">license@1ib-sib.ch</a> ).			
CC				
DR	EMBL; Z46921; CAAB7023.1; -			
DR	PIR; S50358; S50358.			
DR	SGD; S0001431; Y1169C.			
DR	InterPro; IPR004089; Chtax1s transd.			
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2, 1.			
KW	Hypothetical protein; signal.			
FT	SIGNAL 1 23	POTENTIAL.		
FT	CHAIN 1 24 995	HYPOTHETICAL PROTEIN Y1169C.		
FT	DOMAIN 26 253	METHYL-ACCEPTING TRANSDUCER.		
FT	CARBOHYD 28 28	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 35 35	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 468 468	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 664 664	N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	SEQUENCE 995 AA; 99735 MW; F63E287A03F137EC CRC64;			
SC				
Query Match	5.9%; Score 136; DB 1; Length 995;			
Best Local Similarity	17.3%; Pred. No. 3.3;			
Matches	87; Conservative 109; Mismatches 186; Indels 120; Gaps 19.			
QY	18 GTNSS--PSTONTVSRREV-----VSSVQLSEESTFTLCPPVGSFVIRLEFG-----CM 65			
DB	78 GVSSSVGSSSSSVSSVSSGSSSSVSVSSVSSGSS--SSASDVSSVSSGASSSTSDVSS 135			
QY	66 SITKDDANPNNGCQQLERARELTDLINAKAMTTLASLDYAKITEASLSSAYSEAEVTNNN 125			
DB	136 SVSSSSSSASDVSSVSSVSSSSSSASDVSSVSSASASD-----VSSVSSGSSASSTSD 188			
QY	126 LNTLEQLAKMKTNLESALINQANTDKTTFDNEHNVLVYAKLTKTLEQRTNLEGAST 185			
DB	189 VSSSVSSGSSSSASDVSSVSSVSSGSSSSASDVSSVSSGSSASSTSDVSSVSSGASSSTSD 248			
QY	186 AYNQIRNNLVLYNKASLIRKTLDPNLGTLILNSBITTVNRRINNTLSTINEQKTNAD 245			
DB	249 GSGSVSS-----AGSSSSSPFQSTSSASATSSASATSSLSITSSASSASATPSSLSSTSD 304			
QY	246 -----ALNSNPFK-----KVIONNEQSF--VGTITMANVQPSN 276			
DB	305 GTIVLPPTTISGDLITLGGKVIATGEGVVAAGAKLITLIDGGRYSPADLKVAGDILLVKKSK 364			

```

QY 277 YSFVPSADVTVPVNYKYARTVWNGDE---PSSRIILNTNSITDVSMTYSLA-GTNTKYQ 332
D 365 ETPGTEPFDISGENPDTVGN--FNAESBAASTASITSPSSPDNSGDISLSKSKGE 422
QY 333 FSPFSNYPSTGYLYFPYKLVKADANNVGLQVTLNNGNVQVFEATST--SANTTANPT 390
D 423 VTFSPYNSGAFSFGF-----SNAL-----LNGSGVSLGRRDTEGVSNNGNIN-- 465
QY 391 PAVDEIKVAKIVLSGLRFQNTLELSPVPGEGNMKVAFMIGNITY----- 436
D 466 -----LDNSTYIVVPPVSGKGVNIIS---GNLYLHYDPTFTGTVPVK 507
QY 437 -----SSENNADKIP--GY 449
D 508 GEGVLAVDPTETNATITPVGV 529

RESULT 13
ID MSN2 YEAST STANDARD; PRT; 704 AA.
AC P3378;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein MSN2 (Multicopy suppressor of SNF1 protein 2).
GN MSN2 OR YMR037C OR YMR513.02C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
(1)
RN RP SEQUENCE FROM N.A.
RC STRAIN=8288C;
RC MEDLINE=93309420; PubMed=8321194;
RA Estruch F., Carlson M.;
RT "Two homologous zinc finger genes identified by multicopy suppression
RL in a SNF1 protein kinase mutant of Saccharomyces cerevisiae.";
RL Mol. Cell. Biol. 13:3872-3881(1993).
(2)
RN RP SEQUENCE FROM N.A.
RC STRAIN=8288C / AB972;
RC PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Deaman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagals K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RL XIII.";
RL Nature 387:90-93(1997).
(3)
RN RP CHARACTERIZATION.
RP MEDLINE=96208509; PubMed=8641288;
RA Martinez-Pastor M.T., Marchler G., Schueller C., Marchler-Bauer A.,
RA Rius H., Estruch F.;
RT "The Saccharomyces cerevisiae zinc finger proteins Men2p and Men4p
RT are required for transcriptional induction through the stress
RT response element (STRE).";
RT EMBL J. 15:2227-2235(1996).
CC -1- FUNCTION: POSITIVE TRANSCRIPTIONAL FACTOR THAT ACTS AS A COMPONENT
CC OF THE STRESS RESPONSE SYSTEM. RECOGNIZES AND BINDS TO THE
CC STRESS RESPONSE ELEMENT (STRE) WHICH IS INVOLVED IN THE RESPONSE
CC TO VARIOUS FORMS OF STRESS (HEAT, OXIDATIVE, OSMOTIC, ETC.).
CC INVOLVED IN THE REGULATION OF THE CTT1, DDR2, HSP12 GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Contains 2 C2H2-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.1sb-sib.ch/announce/

```

```

CC or send an email to license@1sb-sib.ch).
CC -----
CC DR EMBL; L08838; AAA34806.1; -
CC DR EMBL; Z48502; CAA8403.1; -
CC DR PIR; S39004; S39004.
CC DR HSPD; P08047; ISP1.
CC DR TRANSFAC; T01257; -
CC DR SGD; S0004640; MSN2.
CC DR GO; GO:0005829; C:cytosol; IDA.
CC DR GO; GO:0005634; C:nucleus; IDA.
CC DR GO; GO:0006950; P:response to stress; IMP.
CC DR InterPro; IPR007087; Znf C2H2.
CC DR Pfam; PF00966; Zf-C2H2; 2.
CC DR ProDom; PD000003; Znf C2H2; 1.
CC DR SMART; SM00355; Znf C2H2; 2.
CC DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
CC DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
CC DR DNA-binding; Nuclear protein; Zinc-finger; Metal-binding; Activator;
CC KW Transcription regulation; Repeat.
CC FT DOMAIN 1 109 ASP-RICH (ACIDIC).
CC FT DOMAIN 2 260 279 ASP-RICH (ACIDIC).
CC FT ZN_FING 647 665 C2H2-TYPE 1.
CC FT ZN_FING 676 698 C2H2-TYPE 2.
CC FT ZN_FING 676 698 C2H2-TYPE 2.
CC SQ SEQUENCE 704 AA; 77860 MM; EDF6F07446819DF1 CRC64;

Query Match 5.8%; Score 134.5; DB 1; Length 704;
Best Local Similarity 20.0%; Pred. No. 2.6;
Matches 108; Conservative 83; Mismatches 191; Indels 159; Gaps 26;

QY 15 ILVGTNS-----SPSTQ-----NVTSEVSVQSEESFPYLCPPVGSVIR 59
D 138 ILVMMNPLOSPPSSSPQNPPTNPINTASNETNSPOTSNGNETL-ISPRAQHTSIK 136
QY 60 -----LEFGMSITKQDANPNNGQTOLEARMETDLINAKAMTLASLODYAKIEASLSS 114
D 197 DNRSLPNCANSMLFIDTNPNNINEKL--RNQINSPTNSYSNSISN--SNSNSTGLNS 251
QY 115 AYSAEFVNNNNVANTLEQLQAKTNLESAINQANTDKTTD---NEHPVLVAYALKT 170
D 252 SYFN---SLNDSMLDDVVSDDLNDDDDTNLSRRRSRVITNQFPMTNRSNISH 307
QY 171 TLE-----QRATNLE-GLSTAYNQIRNNLYDLYNKASSLITTKTLDPLNGTILD 219
D 308 SLIDLMNPKINPNRNNTNLTITNTSSSSNASVNTTTNANNADNSNLAG--NPKNDATID 365
QY 220 SNEITTV---NPNINNTLSTINEQKTNADALSFTIKVYQ-----NNEQSF 263
D 366 -NELTOILEMYNNNFNDNLGTSTGK-NKSACPSFEDANAMTKINPSQQLQQQLNRFVQHK 423
QY 264 VGTFTNANVQPSNYSFVAFSADVTVPVNYKYAR-----TYWNGDE--PSSR 307
D 424 QLTSSHNH--SSTNKSFPNSDL-----YSRQRAASLPIDDSLYDLVNRKQDEDPKND 474
QY 308 ILANTNSITDVSMTYSLAGNTKYQPSFNSYGPSTGYLYFPYKLVKADANNVGLQYKL- 366
D 475 MLFNSMLSSQQIK-----PSMIL-----SDNASVIAKYA 505
QY 367 NNGVQVFEATSTSANNTTANPTVAVDEIKVAKIVLSGL----- 406
D 506 TTLSLNDMPFLTEEGON--ANSTPNFD--LSITQNMABLSPASSSSSTSLATNHFYHPP 562
QY 407 RFQNTLELSPVPGEGNMKVAFMIGNITYLSENN-----ADKIPGYR 450
D 563 QQSHHTNMSKIGSLRRKSAVPLMGTVPLTNQNNISSSVNSGTNGAGVTYKRRPSYR 622
QY 451 R 451
D 623 R 623

RESULT 14
PST1_YEAST STANDARD; PRT; 444 AA.
ID PST1_YEAST

```

Query Match 5.8%; Score 134; DB 1; Length 444;  
Best Local Similarity 18.0%; Pred. No. 1.5;  
Matches 89; Conservative 95; Mismatches 185; Indels 126; Gaps 20;

FT	CARBOHYD	263	263	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	268	268	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	280	280	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	292	292	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	305	305	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	329	329	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SO	SEQUENCE	444 AA;	45776 MW;	230F06CACA5921A4	CRC64;

Query Match 5.8%; Score 134; DB 1; Length 444;  
Best Local Similarity 18.0%; Pred. No. 1.5;  
Matches 89; Conservative 95; Mismatches 185; Indels 126; Gaps 20;

QY	1	MHYRRNCIFFLIYLIXGTNS---	PSTQVTSSEVSVSVQSEESTFYLCPPNGSTV	57	
DB	3	LHSLIATALLITTSALATATSSSSIS	PSSCTTISHATATA---QSDLDKYSRCDTLVNL	59	
QY	58	I--FLEFGCMSITKQDANPNNGQ---	TDLEARME-LTDLINAKMTLASLDY	105	
DB	60	IGGLKTKGALANVKE---	INGSLITNATNLISFPADSLSEITDSLNLQSLTILT---	111	
QY	106	AKIBASLSAYSEAEYVNNMLNATLEQ	KKAKTNLESAINQANTDKTTFDNEHNLVEAY	165	
DB	112	-----SASPGSLQSVDSIKLTL	PLPAISFPTSIRKANNIYISDTSL-----QSVDF	158	
QY	166	KALTTLEQRATNIEGLSS--	TANQNRNVLVDLYNKAASLITLDP--	220	
DB	159	SALKRVNPFVNVNKKLTISKSP	EYVSDSLQSFENQTKI--TFPDLVANNISLTV	216	
QY	221	NEITTVNRN-----	INNTLSTINEOKTNA-----DALSNFIRK-----	254	
DB	217	HSVFPALQKINSLSGFINSIS	LNTKLTICGTSTIYSNDLTKLSPSNLSTIGAL	276	
QY	255	VIQNN-----	EQSFVGTFTNAVQPSNTSFVA	SADYTPVNVKART	297
DB	277	VVANNTGLQKIGLDNLTTIG	LTLELVGNFTSLNL--DSLKSXVGADVESKSNFSCNA	334	
QY	298	VWNGDEFSRLIANTNSIT	DVSVYISLAGNTTKYQGFSPN	GPSTGYLTPPYKLVKADA	357
DB	335	L-----	KALQKGGIKGESSFYCKNGAS	TSVKLSSTKSKQSS-----QTTAKVKS	380
QY	358	NNVGLQYKLNNGVQOVPAT	STANNTPPAVDEIKAKIVLSGLRPGNTIELSV	417	
DB	381	SSKKEEKFTSGDKAAASAS	SVSSGAS-----SSSKSKGMAIMA	424	
QY	418	PTGEGNMKVAPMIG	432		
DB	425	PIGQ-----	TTPLVIG	434	

RESULT 15  
YM96\_YEAST STANDARD; PRT; 1140 AA.  
ID\_YM96\_YEAST 004853;  
AC 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Hypothetical 113.1 kDa protein in PRE5-FET4 intergenic region.  
GN YMH317W OR YM99924.09.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C / AB972;  
RX PubMed=9169872;  
RA Bowman R., Churher C.M., Badcock K., Brown D., Chillingworth T.,  
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
RA Jagele K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,  
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrall B.G.,  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
XIII.".  
RL Nature 387:90-93(1997).



```
CC -! DOMAIN: CONTAINS MANY SER/THR-RICH DOMAIN AND REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z54141; CAA90835.1; -.
DR SGD; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Query Match 5.8%; Score 134; DB 1; Length 1140;
Best Local Similarity 21.2%; Pred.No. 5;
Matches 101; Conservative 76; Mismatches 178; Indels 122; Gaps 21;

QY 19 TNSSPSTONVTSREYVSVSVQLESEESTFYLCPPVGVSTVIRLEFGCMGSIKKDANPNNGQ 78
DB 412 TSSSVSSSEISSTKSSVMSSEVSATSLVSSAP--SAISSLASRLPSSK--NTSVTS 466
QY 79 TQLEAARMETDLINAKAMTLASLDYAKIEASLSAYSEAEYNNNNINATLEQLKMAKT 138
DB 467 TLVATEASSVTSLSLPSSETLAS--NSIIESSLSTGYN--STVGSTTSAASSTLGSKVS 521
QY 139 NLESAINQANDTKTFEDNEHPVLVEAYKAKLTLEQRATNL-----EGLSTAYNQ 189
DB 522 SSNSRMATSKTISTSSDLSKSVIFGNSVTYTPSPASISLTASPLPVSWSDTISSEAS 581
QY 190 IRNNLV-----DLVKASSLITKTLDPPLNGTL--LDSEITTVNRNINNTLSTIN 238
DB 582 ISSNLASSAPBDNNSTIASASLIYTKNSVSVSSITSSF--TTNESNLATSTSL 640
QY 239 EOKTNADALNSFIKKVIQNEQSF---VGTFTNANYQPSNYSPAFSADVPVNYKYA 294
DB 641 SNKATARSLSTS-----NATSASNPVPTGTFP-----SMSHTSVITP----- 677
QY 295 RRTVNGDEPPSSRIIANTNSITDVSWIYSLAGTNTKIQSFNSNYGPGSTGYLFPYKLYKA 354
DB 678 -----GFSTSSASLAINSTVVS---SSLAG---YGFSTPESSPTTS-----TLVTS 717
QY 355 ADANNVGLQYKLNNGNVQVEPATSTSA---NNTTANPTPAVDEIKVAKIYLSGLRFGQ 410
DB 718 EAPSTVS-----SMTSAPFINNSTARSPTASAF-----ITR 751
QY 411 NTIEL-SVPTGEGNNK-----VAPMIGNIYLSNENNADKI--PGYRRPG 453
DB 752 STSSISVPLASGDVTSILAANLTFPSAPSTSSAQVLVSKVTSSSILVTPRIDRSG 808
```

Search completed: August 14, 2003, 10:21:39  
Job time : 27 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 14, 2003, 10:18:32 / Search time 101 Seconds

(without alignment)  
1165.069 Million cell updates/sec

Title: US-09-901-572A-3

Perfect score: 2324

Sequence: 1 MHYFRNRKIFLYLYGTN.....SSNNENADKIFGYPRTFL 456

Scoring table: BLOSUM62

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_23.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1207	51.9	271	2	Q49464	Q49464 mycoplasma
2	838	36.1	671	2	Q918D6	Q918D6 mycoplasma
3	828	35.6	702	2	Q49499	Q49499 mycoplasma
4	825	35.5	632	2	Q9XC68	Q9XC68 mycoplasma
5	800	34.4	584	2	Q9KH13	Q9KH13 mycoplasma
6	757	32.6	680	2	Q9KH14	Q9KH14 mycoplasma
7	750	32.3	702	2	Q49498	Q49498 mycoplasma
8	746.5	32.1	645	2	Q921D1	Q921D1 mycoplasma
9	745.5	32.1	644	2	Q92HR9	Q92HR9 mycoplasma
10	745	32.1	650	2	Q49495	Q49495 mycoplasma
11	742.5	31.9	486	2	Q8RLX9	Q8RLX9 mycoplasma
12	741.5	31.9	649	2	Q49497	Q49497 mycoplasma
13	711.5	30.6	647	2	Q49468	Q49468 mycoplasma
14	703.5	30.3	703	2	Q05122	Q05122 mycoplasma
15	703	30.2	656	2	Q9KH15	Q9KH15 mycoplasma
16	618.5	26.6	419	2	Q918D5	Q918D5 mycoplasma

17	579	24.9	386	2	Q49500	Q49500 mycoplasma
18	550.5	23.7	367	2	Q9XC67	Q9XC67 mycoplasma
19	314	13.5	805	12	Q98Y44	Q98Y44 turkey hept
20	314	13.5	805	12	Q98Y45	Q98Y45 turkey hept
21	314	13.5	865	12	Q83291	Q83291 marek disea
22	314	13.5	865	12	Q83JW2	Q83JW2 turkey hept
23	314	13.5	865	12	Q83JW3	Q83JW3 turkey hept
24	314	13.5	865	12	Q83JW4	Q83JW4 turkey hept
25	195.5	8.4	864	12	Q69408	Q69408 melegrid h
26	195.5	8.4	870	12	Q9DP09	Q9DP09 melegrid h
27	195.5	8.4	870	12	Q981G4	Q981G4 melegrid h
28	189.5	8.2	9439	16	Q8CP76	Q8CP76 staphylococ
29	181.5	7.8	865	12	Q69406	Q69406 turkey hept
30	177.5	7.6	865	12	Q9PW21	Q9PW21 turkey hept
31	177.5	7.6	865	12	Q53653	Q53653 staphylococ
32	176.5	7.6	946	16	Q8NXJ1	Q8NXJ1 staphylococ
33	174	7.5	4688	16	Q9EP08	Q9EP08 ureaplasma
34	170.5	7.3	1302	2	Q49547	Q49547 mycoplasma
35	170	7.3	6713	16	Q99U54	Q99U54 staphylococ
36	170	7.3	6713	16	Q99U54	Q99U54 staphylococ
37	170	7.3	6713	16	Q99U54	Q99U54 staphylococ
38	168.5	7.2	3890	16	Q99U53	Q99U53 staphylococ
39	166.5	7.2	3890	16	Q99U53	Q99U53 staphylococ
40	166	7.1	661	16	Q8YMD8	Q8YMD8 anabaena sp
41	164.5	7.1	2481	16	Q99GR6	Q99GR6 staphylococ
42	162.5	7.0	810	5	Q814Y6	Q814Y6 plasmidium
43	162.5	7.0	1237	16	Q9ZK57	Q9ZK57 helicobacte
44	162.5	7.0	1461	16	Q8ZL64	Q8ZL64 salmonella
45	161.5	6.9	1774	5	Q81A08	Q81A08 plasmidium

## ALIGNMENTS

## RESULT 1

ID	Q49464	PRELIMINARY;	PRT;	271 AA.
AC	Q49464;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	TM-1 (Fragment).			
GN	TM-1.			
OS	Mycoplasma gallisepticum.			
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
OX	NCBI_TaxID=2096;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94025893; PubMed=8212828;			
RA	Saito S., Fujisawa A., Ohkawa S., Nishimura N., Abe T., Kodama K.,			
RA	Kamogawa K., Aoyama S., Iritani Y., Hayashi Y.,			
RT	"Cloning and DNA sequence of a 23 kilodalton polypeptide gene of			
RT	Mycoplasma gallisepticum as a possible protective antigen."			
RL	Vaccine 11:1061-1066 (1993).			
DR	EMBL; S65869; AAB28343.2; -			
FT	NON TER 271			
SQ	SEQUENCE 271 AA; 29817 MW; 8825D0CD5C85CA2 CRC64;			

QY	64	CMSTTKKQDANPNQOTQLEAARMELTDLINKAMTILASLDYAKIEASLSAYSEAEVFN	123
DB	27	CMSTTKKQDANPNQOTQLEAARMELTDLINKAMTILASLDYAKIEASLSAYSEAEVFN	86
QY	124	NNLWATLEQOLMAKTNLSAIAINQANTDKTTFDNEHPNIVEAYKALKTTLEORATNLEGS	183
DB	87	NNLWATLEQOLMAKTNLSAIAINQANTDKTTFDNEHPNIVEAYKALKTTLEORATNLEGS	146
QY	184	STAYNOIRNNLVLDYNNASSLITKTLDPPLNGTLLDSNEITTVNRNINNTISTINEQKTN	243
DB	147	STAYNOIRNNLVLDYNNASSLITKTLDPPLNGTLLDSNEITTVNRNINNTISTINEQKTN	206

```

OY 244 ADALNSPFIKKYIIONNEOSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTVWNGDE 303
DB 207 ADALNSPFIKKYIIONNEOSFVGTFTNANVOPSNYSFVAFSADVTVPVNYKARRTVWNGDE 266
OY 304 PSSRI 308
DB 267 PSSRI 271

RESULT 2
OY 09L8D6 PRELIMINARY; PRT; 671 AA.
AC 09L8D6;
ID 09L8D6;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE PMGA-like protein 9.2.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F,
RA Pharr G.T., Branton S.L., Hanson L.A., Minion F.C., Lott B.D.,
RT May J.D., Hugliet M.B.,
RT "A novel PMGA-like gene from the F-strain (vaccine strain) of
RT Mycoplasma gallisepticum."
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF210770; AAF29524.1; -
SQ SEQUENCE 671 AA; 71898 MW; 4DA8E6591750679 CRC64;

Query Match
Best Local Similarity 40.7%; Pred. No. 2.3e-32;
Matches 184; Conservative 78; Mismatches 128; Indels 62; Gaps 10;

OY 53 VGSYTRLEFGCMSTTKDANP-----NNGOT-----OLENA 84
DB 15 ISSFWLAAASCTSATPTTPPEPKPDPMPNPPSGDNGDTPNPGDGMENSAOQLAA 74
OY 85 RMELDLINAKAMTLASLDQYAKIEASLSAYSEAEVNNNNLNATLEOLKMAKTULESAI 144
DB 75 KKELDLATQSNLAKADYNTNIGTLTAAYTTAKSTSDNTSVLEQVKSATSTLQAAI 134
OY 145 NOANTDKTTFDNEHPLVAVYALKTTLEQKATNIEGLSTVNOIRNVLDLYNKASL 204
DB 135 DTAASKTISFDEKNPELIKAYVALKELTKNEEIVLSGLTDSNPFATIKTNLTALYOSGDF 194
OY 205 IRTKLDLNGGTLTLDSEITTVNRNINNTLSTINEQKNADALNSPFIKKYIIONNEOSFV 264
DB 195 VKATLDPVSGNA-POIADITKADKDIADAVSKLETWKTNANTLATSFEVAVKKTLCI 253
OY 265 GFTTNANVOPSNYSFVAFSADVTVP-----VNYKARRTVWNGDEPSSRIIA-----NTNSIT 316
DB 254 DT-TNNREQPGVNSFVGYSVNATNNNEIPIWNPFAQKWTSDNGRTSLISTSDMSITLT 312
OY 317 DVSMTYSLAGTNTKYOPSSSNVGPSTGYLYFPYKLVKADANNVGYOVLNNGNVQOAVE 376
DB 313 EVSWIYSLGAGTKYSLTFTNYGPGSTGYLYFPYKLVKADANNVGYOVLNNGNVQOAVE 372
OY 377 A-----TSTA-----NNT-----ANPTPAVDEIKVAKIVLSGLRFGONTIE 414
DB 373 APVTKTSVADSGDSDNNQTESAAETMPVTSIDNPATVSDINIAITLTSNLFSGSTIE 432
OY 415 LSVPTGEGNNKVAIPMIGNTYSSNNADKI 446
DB 433 FSVPTEPS--NKVAPMIGMYLTSTNIAEAKV 462

RESULT 3
OY 094999 PRELIMINARY; PRT; 702 AA.
AC 094999;
ID 094999;

```

```

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PMGA1.4 protein precursor.
GN PMGA1.4.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56;
RX MEDLINE=95010739; PubMed=7925999;
RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,
RA Browning G.F., Whitehead K.G., Walker I.D.;
RT "The organisation of the multigene family which encodes the major cell
RT surface protein, PMGA, of Mycoplasma gallisepticum."
RL FEBS Lett. 352:347-352(1994).
DR EMBL: L28424; AAA62418.1; -.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 25 POTENTIAL.
B70AC874FE85055C CRC64;

Query Match
Best Local Similarity 40.6%; Pred. No. 7.3e-32;
Matches 176; Conservative 86; Mismatches 125; Indels 46; Gaps 9;

OY 50 PPPGVSIVILFEGCMSTTKDANPNNG-----QTOLEAARMELTDLINAKAMTLASLQ 103
DB 54 PNPFGGMGNGNGG-----NTNPGNGGGTNDNAOQLAAKKELSLTLQNSLSTYA 106
OY 104 DYAKIEASLSAYSEAEVNNNNLNATLEOLKMAKTULESAI NOANTDKTTFDNEHPLV 163
DB 107 DYANIGTLTAAYTTAKSTSDNTSATLEQVKSATSTLQTAIDTAASKTSFDEKNPELIK 166
OY 164 AYKALKTTLEORATNIEGLSTVNOIRNVLDLYNKASLITLTDPLNGGTLTLDSEI 223
DB 167 AYNAIKETLKKMRNSLGLTDSNPFATIKTNLTALYOSGKDVITLDPIM-GTAINISAV 225
OY 224 TVVRNINNTLSTINEQKNADALNSPFIKKYIIONNEOSFVGTFTNANVOPSNYSFVAFS 283
DB 226 SOANTNINSAVSKLETWKTNANTLATSFEVAVKKTLCI-DT-TNNOEQPGVNSFVGY 284
OY 284 ADVTP-----VNYKARRTVWNGD-----EPSSRIIANTNSITDVSITSLAGTNTKY 332
DB 285 VDVTTGSDNAPPNMSPQKRWTSNTDILSQPQAEGENQOSADVSIVYVLTGMAKYS 344
OY 333 FSPSNVGSSTGYLYFPYKLVKADANNVGYOVLNNGNVQOAVEATS-----TSAN 383
DB 345 LTFNYGPGSTGLFPYKLVNSSDKVALLEYKLNESA VKTIDSPSGTSPVADATREN 404
OY 384 N-----TTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAIPMIG 432
DB 405 NRSTAPAPQGSTINPAATLTDIKAKVTLNLFSGSTIEFSVPTAKEGTSKVAIPMIG 464
OY 433 NIYSSNNADK 445
DB 465 NMVLTSSDRDYK 477

RESULT 4
OY 09XCG8 PRELIMINARY; PRT; 632 AA.
AC 09XCG8;
ID 09XCG8;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE VHA1 precursor (Fragment).
GN VHA1.
OS Mycoplasma imitans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
NCBI_TaxID=29560;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=4229; PubMed=10463176;  
 RA MEDLINE=99392472; Duffy M.F., Glew M.D., Browning G.F.;  
 RT "A gene family in Mycoplasma imitans closely related to the pmcA  
 RU Microbiology 145:2095-2103(1999).  
 DR EMBL: AF141940; AAD39463.1; -.  
 KM Signal.  
 FT NON\_TER 1 17 POTENTIAL.  
 FT SIGNAL 18 632 VLHAI  
 FT CHAIN 18 632  
 SQ SEQUENCE 632 AA; 66959 MW; 173FSB12E705BE47 CRC64;  
 Query Match 35.5%; Score 825; DB 2; Length 632;  
 Best Local Similarity 43.4%; Pred. No. 8.9e-32;  
 Matches 187; Conservative 68; Mismatches 130; Indels 46; Gaps 11;

QY 53 VGSIVIRLEFGC-MSITKGDANPNN-----GQT---QLEARMELTDLIN 93  
 Db 7 IGSFVTLAAASCSKAVTSNPGNPENPKPNPDGTGNTPGDTPAKQQLADAKASLNTLLG 66  
 QY 94 AKAMTLASLODYAKIEASISAYSEAEFTVNNNLNATLEQLKMAKTNLESAINQANTDKT 153  
 Db 67 SQSTNVVALYEDYAKIKOTLSAIVASAOQTANNANATLKVDADAKAQLATSDANAKTE 126  
 QY 154 PDNEHPNLVEAYKALKTLEQATNLEGLSSTAVNQIRNNLVLYNKASSLITKLDPLN 213  
 Db 127 FDKANLGLVSAVAKETLKSETNLDGLSODNYSALIANVSLYNKAKDTTYTLDP-T 185  
 QY 214 GGTLLDSNEITTVNINNTLSTINEOKTNADALSNSFIKVIQNNESFVGTFTN---- 269  
 Db 186 SGMIPTKDEITSANTALTOAVSAIDSOQTNADTAATFEIKELDSAKLT--GTTAEQQA 244  
 QY 270 ANVQPSVYFVAFSADYTP-----VYKKYARRYVNNGDPS--SRLANTNSIT 316  
 Db 245 ASQPGNSFVGFSDVDVTGTGSGOEDLPFWMNPAKRKVMTEGLSACQTLVSSSETPLT 304  
 QY 317 DVSMTYSLAGTNTKYQSFNSNYGPGSTGYLYFPYKLVKADANNVGLQYKLNNG--NVOQVE 375  
 Db 305 DVSMTYSLTGAGSKYTLFTFYGPSTALVLPYKLVQSSDNNKLGLOKLANSETLVPIT 364  
 QY 376 FATSTANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAEMIGNIY 435  
 Db 365 FGNFTN----DSGATPAIDIKVEKVTLSINFGDNTIEFSVATEE--NKVAPMIGMY 417  
 QY 436 LSSNENNAKDI 446  
 Db 418 LTTSTNNVDKI 428

RESULT 5  
 Q9KH13 PRELIMINARY; PRT; 584 AA.  
 ID Q9KH13  
 AC Q9KH13  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 RT "Sequence analysis of the pmcA multigene family of Mycoplasma  
 RL gallisepticum strain HS."  
 DE Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DE EMBL: AF275312; AAF91415.1; -.  
 DE InterPro: IPR001986; EPPS\_synase.  
 DR PROSITE; PS0104; EPPS\_SYNTHASE\_1; 1.  
 FT NON\_TER 584 584

SQ SEQUENCE 584 AA; 62453 MW; 5C467BA55FB27A72 CRC64;  
 Query Match 34.4%; Score 800; DB 2; Length 584;  
 Best Local Similarity 42.1%; Pred. No. 1.3e-30;  
 Matches 179; Conservative 71; Mismatches 115; Indels 60; Gaps 10;

QY 73 NPNNGQT---QLEARMELTDLINAKAMTLASLODYAKIEASISAYSEAEFTVNNNLN 128  
 Db 69 NPNNGTTPBQQLAAAKKTLTDLTGENTNVALYADYAKIQSTLSTAYMTAKTASENTSA 128  
 QY 129 TLEQLAKNTLSAINQANTDKTTPNEHPNLVEAYKALKTLEQATNLEGLSSTAVN 188  
 Db 129 TLDNRASSTTLQALIDKASNKRTEFSAQPLVATVNOUKTTLQSTKTSLEGLSEKYS 188  
 QY 189 QIRNMLDLNKNKASLITKTLDPINGTLLDSNEITTVNINNTLSTINEOKTNADALS 248  
 Db 189 SIKNHLSPDAGSALAKTLDP-TMGITVEVNSVTANEDIMTAVSKLTETKTNADKF- 246  
 QY 249 NSFPIKVIQNNESFVGTFTNAN--VQPSNYFVAFSADYTP-----VYKKYART 297  
 Db 247 NDFEKKPL--SKELVETNDRAHNQEPANMSFAGYVDLTGSGTNSQNLPMNPFQPK 304  
 QY 298 VVNGD---EPPSRLANTNSITDVSMTYSLAGTNTKYQSFNSNYGPGSTGYLYFPYKLVKA 354  
 Db 305 VMTSEGOQTGKTLVSSPVSAIDVSWIYSLAGETKYTLSEFYYPGPTAFLYFPYKLVKQ 364  
 QY 355 ADANNVGLQYKLNNGNQOVEF-----ATST-----S 381  
 Db 365 ADSSVALQYSLNKTSSKLINFPAKTMPTNADOSEGVATTSTTEGRSSSEVLVADEVA 424  
 QY 382 ANNTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNNKVAEMIGNIYLSNEN 441  
 Db 425 AVNNENMPTPLVSINIAKVTLSGLTGEENTIEFSVPT-----NKVAPMIGMYLTSNSG 479  
 QY 442 NADKI 446  
 Db 480 SGGKI 484

RESULT 6  
 Q9KH14 PRELIMINARY; PRT; 680 AA.  
 ID Q9KH14  
 AC Q9KH14  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 RT "Sequence analysis of the pmcA multigene family of Mycoplasma  
 RL gallisepticum strain HS."  
 DE Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DE EMBL: AF275312; AAF91414.1; -.  
 DE InterPro: IPR001986; EPPS\_synase.  
 DR PROSITE; PS0104; EPPS\_SYNTHASE\_1; 1.  
 FT NON\_TER 680 680

Query Match 32.6%; Score 757; DB 2; Length 680;  
 Best Local Similarity 39.6%; Pred. No. 1.7e-28;  
 Matches 163; Conservative 71; Mismatches 130; Indels 48; Gaps 7;

QY 71 DANPNNGQTOLEARMELTDLINAKAMTLASLODYAKIEASISAYSEAEFTV 123  
 Db 63 DIVEGGQGMNATNGLVNAKKRALSGLIGESKTVELVADYAKIKADLTSAVVAKTTS 122  
 QY 124 NNLNATLEQLKMAKTNLESAINQANTDKTTPNEHPNLVEAYKALKTLEQATNLEGLS 183  
 Db 123 DSSTSTLDQVKTATNTLQTAINTAASDKERPDQNSQLLMAVYKLVKOTLNKKEAIVNSLN 182

QY 184 STAYNOIRNNLVLDYKNSASLITKTLDPNGTLLDSNEITTVNRNINNTL-----ST 236  
 DB 183 QEKYSALISEINAASTAEIYKOTLVNNG-----NLPVAAALMENTKILEAIKEK 236  
 QY 237 INEKTNAALASNFIFIKVIONNEQSPFGFTNANVOPSNYPFAFSADVPVNYKYARR 296  
 DB 237 INSEKSNADLEFANQYKTL--DRTKLMSEGSNNITKQPGNYSFAYASDIASPPMNRQR 293  
 QY 297 TVNNGD-----EPSSRIANTNSITDVSWIYSLAGTNTKYOFSNNGPSTGYLYPPYKL 351  
 DB 294 TVMTADSRWTSPLEPNLQNGAPLTDVSMIYTLSTGAKYTLTFDYGPQGTGYLYPPYKL 353  
 QY 352 VKAADANNVGYOKLNNNGVQVEF-----ATSSANTTANPTTAPV 393  
 DB 354 VKTSD--KVGLOQYKLNQADPVALIOFSEAAATASAPAEITDGOESAEATANEKVNPMPSV 411  
 QY 394 DEIVAKYIVLSGRFGONTIELSVPTGNNKRYAPMIGNIYLSNENNAKD 445  
 DB 412 NTINAVAKTSLNKGSTNITIEFSVPMQDNMKKAPMIGNIYITSNDEANK 463

## RESULT 7

Q49498 PRELIMINARY; PRT; 702 AA.  
 AC Q49498:  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE PMGAL.3 protein precursor.  
 GN PMGAL.3  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NC NCB1\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S6;  
 RX MEDLINE=95010739; PubMed=7925999;  
 RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
 RA Browning G.F., Whithear K.G., Walker I.D.,  
 RT "The organisation of the multigene family which encodes the major cell  
 RT surface protein, PMGA, of Mycoplasma gallisepticum.";  
 RL FEBS Lett. 352:347-352(1994).  
 DR EMBL; L28424; AAA62417.1; -.  
 KM Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 SQ SEQUENCE 702 AA; 75537 MW; 273E8915FEES7B9F CRC64;

Query Match 32.3%; Score 750; DB 2; Length 702;  
 Best Local Similarity 41.4%; Pred. No. 3.7e-28;  
 Matches 175; Conservative 62; Mismatches 128; Indels 58; Gaps 11;

QY 73 NENNGQT--OLEAARMLTDLINAKAMTASLQDYAKIEASLSANSEATTNNNINA 128  
 DB 69 NENPGNTPEQQLAARKTLTDLTGTEENTVALYDVKIOSTLTAVYATATTAENISA 128  
 QY 129 TLEQKMAKTNLESAINQANTDKTFPNDNEHPLVEAYKALKTTLBOBATNLEGLSTAYN 188  
 DB 129 TLENRASATTLQALIDKANDKRVFDSVNOPLVAAYNNKLTLLSKTSLSGLESENKYG 188  
 QY 189 QIRNNLVLDYKNSASLITKTLDPNG--GTLDSNEITTVNRNINNTLS--TINEOKTNA 244  
 DB 189 GIKNHLSKLPDGSATLAKTLDPTSGERTLEKVN--ANNGIKMAISPELSKKWKGNA 245  
 QY 245 DALNSFIKVIQNNQSFVGTFTNANVOPSNYSFPAFSADVP-----VNTKYARRTW 299  
 DB 246 DLF--NEFEKQPLSKKELKSTSDTAHQEQEPANWSFPAVSVDLTLSNQNLPNNNPAQRKW 304  
 QY 300 NQD--EPPSRILANTN--SITDVSWIYSLAGTNTKYOFSNNGPSTGYLYPPYKLVKAD 356  
 DB 305 TSENQPGKTAIVSSVSAITDVSWIYSLAGEGTKYTLTFEYGPQDAFLYLYKLVKAD 364  
 QY 357 ANNNGLOQYKLNNGVQVEF-----ATST-----SAN 383

DB 365 SSSVALQYSLNKTSSKLNPKPAETVSTNTDSENEVATSTTEARSSYKLVADAENAS 424  
 QY 384 NTTANPTPAVDEIKVAKYIVLSGRFGONTIELSVPTGEGNNKRYAPMIGNIYLSNENNA 443  
 DB 425 NNMENHTPTVDINIAKVTLSGLTFGENTIEFSVPEG-----KVAPMIGNIYLSNSESQ 479  
 QY 444 DK1 446  
 DB 480 VK1 482

## RESULT 8

Q92ZD1 PRELIMINARY; PRT; 645 AA.  
 AC Q92ZD1:  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE M9 protein.  
 GN M9  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NC NCB1\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PG31;  
 RX MEDLINE=99003182; PubMed=9784576;  
 RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.,  
 RT "A protein (M9) associated with monoclonal antibody-mediated  
 RT agglutination of Mycoplasma gallisepticum is a member of the PMGA  
 RT family.";  
 RL Infect. Immun. 66:5570-5575(1998).  
 DR EMBL; AF032890; AAC69269.1; -.  
 SQ SEQUENCE 645 AA; 69876 MW; 5A2F29572CFAE055 CRC64;

Query Match 32.1%; Score 746.5; DB 2; Length 645;  
 Best Local Similarity 40.4%; Pred. No. 4.9e-28;  
 Matches 176; Conservative 67; Mismatches 134; Indels 59; Gaps 12;

QY 53 VGSIVILLEGCGMSITKK-----DANPNNGQ-----TOLEAARMLTDLINA 94  
 DB 15 IGSFVMLAASCTTPTEPNPDPSCGNNQGTNPBGQGMNNAASQELAAARMGLTTFDS 74  
 QY 95 KAMTASLQDYAKIEASLSANSEATTNNNINATLEQLKMAKTNLESAINQANTDKTF 154  
 DB 75 KAKVLGYVDYKKYQDTLTAAYDAKTVLDNNSSTTONLINEAKTRLETAIRTAATSKOTF 134  
 QY 155 DNEHPNVEAYKALKTTLBOBATNLEGLSTAYNQIRNNLVLDYKNSASLITKTLDPNG 214  
 DB 135 DEQHAELVYKVEELKTLTSMETATLAPYAAQVAGIKHLSGLYDAGKALITKLEPVEG 194  
 QY 215 GTLDSNEITTVNRNINNTL--STINEOKTVAALNSFIKVIQNNQSFVGTFTNANV 272  
 DB 195 DP--LTDADVMMANKIYEAIDDEVLPQKEKATLDAISFVKQYLVKEKITVEEAAHN--KA 252  
 QY 273 QPSNYSFPAFSADVPV-----NRYKARRTW--NGDEPSSILANT-----NSI 315  
 DB 253 QPANYSFVGYVDITGVGTQTSIPNDYDQRIITFNGDEP--RSISNTPADGQTMVQPL 310  
 QY 316 TDVSWIYSLAGTNTKYOFSNNGPSTGYLYPPYKLVKADANNVGYOKLNNNGVQVE 375  
 DB 311 SNVSWIYSLAGTNGAKYTLTFEYGPSTGYLYPPYKLVNTSOMKGLGLEYKLAND----- 363  
 QY 376 FATSTSA-----NNTTANPTPAVDEIKVAKYIVLSGRFGONTIELSVPTGEGNNKRYAPM 430  
 DB 364 -ATEPSALTTPSEQTMNCKTPTVDINAVAKTIANLNKRGSKIKFISVDA-----EKVSPM 417  
 QY 431 IGNIYLSNENNAKD1 446  
 DB 418 IGNMILSSSPNNMNKI 433

## RESULT 9

Q9ZHR9 PRELIMINARY; PRT; 644 AA.  
 AC Q9ZHR9; TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE Hypothetical 69.9 kDa protein.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PG31;  
 RX MEDLINE=99003182; PubMed=9784576;  
 RA Liu L., Payne D.M., van Santen V.L., Dybvig K., Panangala V.S.;  
 RT "A protein (M9) associated with monoclonal antibody-mediated  
 RT agglutination of Mycoplasma gallisepticum is a member of the pmcA  
 RT family.";  
 RL Infect. Immun. 66:5570-5575 (1998).  
 DR EMBL; AF053978; AAC69274.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 644 AA; 69866 MW; 8B9F352B13FBDE5C CRC64;

Query Match 32.1%; Score 745.5; DB 2; Length 644;  
 Best Local Similarity 40.4%; Pred. No. 5.5e-28;  
 Matches 176; Conservative 68; Mismatches 133; Indels 59; Gaps 12;

QY 53 VGSIVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTDLINA 94  
 DB 15 IGSFVMLAASCTPTSPAPNPSPGNGGMDTNPBGQGMNNAASQELAAARMGLTVFDS 74  
 QY 95 KAMTLASLQDYAKIEASLSAYSEAEVNNNNLNATLEQLKMAKTNLESAINQANTDKTF 154  
 DB 75 KAKNGLGVVDYKTKQDILTAKYDAKTYLDNSSSTTONLEAKRTLEIRTAITSKOTF 134  
 QY 155 DNEHNLVEAYKALKTTLEQORATNLEGLSTAYNOIRNNLVLDLYNKASLLITKLDPLNG 214  
 DB 135 DEQHAELVKVYBELKTTLSNETATLAPYAAQYAGIKHLSGLYDAGKAITTKLEPVEG 194  
 QY 215 GTLDSNITTVNRRINNTL--STINEOKTNADALSNSFIKVIQNNOSFVGFTTANV 272  
 DB 195 DP-LTASAVMMANTKIVEAIKDEVLNPKKENATKLADSFVKQVLVEKKTGVEAHNKS- 252  
 QY 273 QPSNVSFAVFSADVTPV-----NYKARRTV--NGDEPSRIANT-----NSI 315  
 DB 253 QPAYSFVGYVDITGTANGQTSIPNNMYAQRTIFITNDEP--RVSNTTPVDGOTMAQPL 310  
 QY 316 TDVSWIYSLAGTNTKYQFSNYPSTGYLYPYKLVKADANNVGLQYKLNNGVQVE 375  
 DB 311 SNVSWIYSLAGTNGKTYLFTYGGSTGYLYFPYKLVNTSDQVKLGLEYKIND----- 363  
 QY 376 FATSTSA-----NNTTAPTPAVDEIKYAKIVLSGLRFGONTIELSVPTGSGNNKVP 430  
 DB 364 -ATKPSALTTFGSEQTMNGKTPVNDINAVAKVLANLNGSKNIEFSPV-----VEKSP 417  
 QY 431 IGNITLSSNNENADKI 446  
 DB 418 IGMNVLSSSPNNMKI 433

RESULT 10  
 Q9495 PRELIMINARY; PRT; 650 AA.  
 AC Q9495; O08060;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
 DE Haemagglutinin precursor.  
 GN PMGAL.1.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NCBI\_TaxID=2096;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=S6;  
 RX MEDLINE=95010739; PubMed=7925999;  
 RA Markham P.F., Glew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
 RA Browning G.F., Whitehead K.G., Walker I.D.;  
 RT "The organisation of the multicene family which encodes the major cell  
 RT surface protein, PMGA, of Mycoplasma gallisepticum.";  
 RL FEBS Lett. 352:347-352 (1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S6;  
 RA Markham P.F.;  
 RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; U90714; AAB50152.1; -.  
 KW Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 650  
 FT SIGNAL HAEMAGGLUTININ.  
 SQ SEQUENCE 650 AA; 70249 MW; 3ABACB65940EBB CRC64;

Query Match 32.1%; Score 745; DB 2; Length 650;  
 Best Local Similarity 40.5%; Pred. No. 5.9e-28;  
 Matches 177; Conservative 65; Mismatches 139; Indels 56; Gaps 12;

QY 53 VGSIVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTDLINA 94  
 DB 15 IGSFVMLAASCTPTSPAPNPSPGNGGMDTNPBGQGMNNAASQELAAARMGLTVFDS 74  
 QY 90 DLINAKAMTLASLQDYAKIEASLSAYSEAEVNNNNLNATLEQLKMAKTNLESAINQANT 149  
 DB 75 TVPDSKKNLGLVYDYKTKQDILTAKYDAKTYLDNSSSTTONLEAKRTLEIRTAITSKOTF 134  
 QY 150 DKTFDNEHNLVEAYKALKTTLEQORATNLEGLSTAYNOIRNNLVLDLYNKASLLITKLDPLNG 209  
 DB 135 SKQTFDQHAELVKVYBELKTTLSNETATLAPYAAQYAGIKHLSGLYDAGKAITTKLEPVEG 194  
 QY 210 DPLNGTGLDSNITTVNRRINNTL--STINEOKTNADALSNSFIKVIQNNOSFVGFTTANV 267  
 DB 195 EPGVGD-LTAAVMTANTKIVEAIKDEVLNPKKENATKLADSFVKQVLVEKKTGVEA 253  
 QY 268 TNAVQPSNVSFAVFSADVTPV-----NYKARRTV--NGDEPSRIANT----- 312  
 DB 254 HN-KAOPANSFVGYVDITGTANGQTSIPNNMYAQRTIFITNDEP--RVSNTTPVDGOT 310  
 QY 313 --NSITVSWIYSLAGTNTKYQFSNYPSTGYLYPYKLVKADANNVGLQYKLNNGVQVE 370  
 DB 311 MAQPLSNVSWIYSLAGTNGKTYLFTYGGSTGYLYFPYKLVNTSDQVKLGLEYKINDA- 369  
 QY 371 VOQVEFATSTSAANTTAN-PTPAVDEIKYAKIVLSGLRFGONTIELSVPTGSGNNKVP 429  
 DB 370 ---TEPSAITFGNQTMNGKTPVNDINAVAKVLANLNGSKNIEFSPV-----EKVSP 421  
 QY 430 MIGNITLSSNNENADKI 446  
 DB 422 MIGNVLSSSPNNMKI 438

RESULT 11  
 Q8RLX9 PRELIMINARY; PRT; 486 AA.  
 AC Q8RLX9;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hemagglutinin (Fragment).  
 GN VHA.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=tc-11;  
 RA Markham P.F., Kanci A., Caifra G., Sundquist B., Haines P.,  
 RA Browning G.F.;

"Malt homolog in Mycoplasma gallisepticum is not essential in vitro."

RT Submitted (JDC-2001) to the EMBL/GenBank/DBJ databases.  
 FT EMBL: AY055985; AAL58980.1; -  
 FT NON\_TER 486 486  
 SO SEQUENCE 486 AA; 51844 MW; 94912DD7A09FB911 CRC64;

Query Match 31.9%; Score 742.5; DB 2; Length 486;  
 Best Local Similarity 38.0%; Pred. No. 5.5e-28;  
 Matches 175; Conservative 68; Mismatches 128; Indels 89; Gaps 12;

51 PVGVTIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTDLINAAMTLASIOD 104  
 52 PPSGG-----NNMGSDTNPENGGGMNANVOQLAAKATLTLTLNGTEKVLND 101  
 105 YAKTISLSSAYSEAEYNNNNLNATLEQIKAKNTLESAINQANTDKTFDNEHPNVEA 164  
 102 YAKTIDDLVYAKIAKEISDKSHATLOEVNNAKRTLETAIKDAANSKTSFEKPELKA 161  
 165 YKAKTTLLEQRTNLEGLSTRAYNOIRNNVDLYNKASLITKTLDPINGTLDNSBIT 224  
 162 YDALKQTITSEMSINQMDANFETIKHISNLYKQGDIIITATLDPITG---DGPAM 217  
 225 TYNRN---INNTLTINEOKTNADALSNSFIKVIQN---NEQSFVGTFTNANVQPSNY 277  
 218 VNQNTNEALVNAKTSKIEDKTNATLAFVQKTLNANLVN-----TNNQPSQSY 270  
 278 SFVAESADV-----TPVNYKYARTVW---NGDEP---SSRIANTNSITDVSYTY 322  
 271 SFVAYSVDLNTGVSTASNTF-NMNLAQKRVWVSGSGRTSPSSSDANNSPALTDVSWY 329  
 323 SLAAGNTTKQSFNSYNGPSTGYLYFPYKLYKADANNVGLQYKLNANVOVEFA----- 377  
 330 NLSGANSKYTLTFNMGVSTGHLFPPYKLYKSDQNVGLQYTLNNKPAQRIEFPAPASP 389  
 378 -----TSTANNTANPTPAVDEIKYAKIVISGL 406  
 390 SSGGTRHADPSPRAAATETDVSDSABSGQATQTMSSSMNKTFVSDINAVSTLSDL 449  
 407 RFGQNTIELSVPTGEGNNKVAIPMIGNIYLSNENNAKDI 446  
 450 NEGANTIEFSVPMGD---SMVAPMIGNMYITSNPLNANOI 486

## RESULT 12

Q49497 PRELIMINARY; PRT; 649 AA.  
 ID Q49497;  
 AC Q49497;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE PMGAL.2 protein precursor.  
 GN PMGAL.2  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S6;  
 RX MEDLINE=95010739; PubMed=7925999;  
 RA Markham P.F., Giew M.D., Sykes J.E., Bowden T.R., Pollocks T.D.,  
 RT "The organization of the pmga, of Mycoplasma gallisepticum,"  
 RT surface protein, pmga, of Mycoplasma gallisepticum,"  
 RL FBS Lett. 352:347-352(1994).  
 DR EMBL; L28424; AAA62416.1; -  
 KM Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 SO SEQUENCE 649 AA; 70205 MW; 356554BD2C72C1F8 CRC64;

Query Match 31.9%; Score 741.5; DB 2; Length 649;  
 Best Local Similarity 40.0%; Pred. No. 8.6e-28;  
 Matches 176; Conservative 65; Mismatches 136; Indels 63; Gaps 12;

QY 53 VGSTVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTD 90  
 DB 15 IGSFVMLAASCTTPTNPPTNPENPPSGGMNGDTNPGGQGMNAAASQELAARVGLTT 74  
 QY 91 LINAAMTLASIODYAKIEASISAYSEAEYNNNNLNATLEQIKAKNTLESAINQANTD 150  
 DB 75 IPDSKAKNIGLYVDYKQTQNTLTAKADAKTVLDNSSSTTONLINEAKRTLETAIRTAAS 134  
 QY 151 KTFDNEHPNVEAYKALKTTLLEQRTNLEGLSTRAYNOIRNNVDLYNKASLITKTLTD 210  
 DB 135 KQTFDQHAELVYVELKTLTSLNETATLAPYADAOYAGIKHLSLVSAGAAITKTLLE 194  
 QY 211 PLNGGTLDNSNETTYNRRNNTLT---STINEOKTNADALSNSFIKVIQNNQSFVGT 268  
 DB 195 PVGSDP-LTASAAAMANTKIVEAIKQEVNLPQENATKLADSFVKQVLEKATGVEBAH 253  
 QY 269 NANVQPSNYSFVAFSADVTPV-----NYKARTVW---NGDEPSSRIANT 312  
 DB 254 N-KAQPANYSFVGVSDITGTTGQTSIPNWDYAOHTITNSDEP---RSISNTPADGQTM 310  
 QY 313 -NSITVSWIYSLAGTNTKYQSFNSYNGPSTGYLYFPYKLYKADANNVGLQYKLNANV 371  
 DB 311 AQLPSNVSWIYSLAGTAKTTLFTYGGSTGYLYFPYKLYKSDQNVGLQYKLEKLDND 367  
 QY 372 QOVEPATSTSA-----NNTTANPTPAVDEIKYAKIVISGLRFGQNTIELSVPTGEGNNK 426  
 DB 368 -----ATKPSAITFGSDQTNNGKTFVNDINAVKTLANLNFSGNKEFSVPA-----EK 417  
 QY 427 VAPMIGNIYLSNENNAKDI 446  
 DB 418 VSPMIGNMYLSSPNNMKI 437

## RESULT 13

Q49468 PRELIMINARY; PRT; 647 AA.  
 ID Q49468;  
 AC Q49468; Q53303;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hemagglutinin homolog precursor.  
 GN PMGAL.2  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=9162830; PubMed=8432610;  
 RA Markham P.F., Giew M.D., Whitehead K.G., Walker I.D.,  
 RT "Molecular cloning of a member of the gene family that encodes pmga, a  
 RT hemagglutinin of Mycoplasma gallisepticum,"  
 RL Infect. Immun. 61:903-909(1993).  
 DR EMBL; M83178; AAA02996.1; -  
 DR EMBL; S55216; AAB25397.2; -  
 KM Signal.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 647 HEMAGGLUTININ HOMOLOG.  
 SO SEQUENCE 647 AA; 70333 MW; 33916673B89528C4 CRC64;

Query Match 30.6%; Score 711.5; DB 2; Length 647;  
 Best Local Similarity 39.0%; Pred. No. 2.3e-26;  
 Matches 173; Conservative 63; Mismatches 135; Indels 73; Gaps 13;

QY 53 VGSTVIRLEFGCMSTTKK-----DANPNNGQ-----TQLEAARMELTD 90  
 DB 15 IGSFVMLAASCTTPTNPPTNPENPPSGGMNGDTNPGGQGMNAAASQELAARVGLTT 74  
 QY 91 LINAAMTLASIODYAKIEASISAYSEAEYNNNNLNATLEQIKAKNTLESAINQANTD 150  
 DB 75 IPDSKAKNIGLYVDYKQTQNTLTAKYDAKTVLDNSSSTTONLINEAKRTLETAIRTAAS 134  
 QY 151 KTFDNEHPNVEAYKALKTTLLEQRTNLEGLSTRAYNOIRNNVDLYNKASLITKTLTD 210

Db 135 KOTPEOHAELVKYKELKTLTSLNETATLAPADAQVAGIKMHLGSLYDACKAITTKLE 194  
 QY 211 PLNGGTLDSNEITVNNINNTL--STINEOKTN---ADALSNFICKVYIONNEQSFV 264  
 Db 195 PVEBDP-LTASAVMANAKTYEALKEDEVLNPOKENTATLADSLSSLYKKTGVEE---- 249  
 QY 265 GTFTNANVOPSNYSFVAFSADVTVP-----NYKVARFTW-NGDEPSSRIIANT--- 312  
 Db 250 ---AHNKQAPANYSFVGYKRWYTELDDKQVFPNMDVQORTIFTNSEDP--RSISNTPAD 304  
 QY 313 -----NSTDYSWYSLAGTNTKYQFSPSNYGPSTGYLYFPYKLVKADANNVGLQYTLN 367  
 Db 305 GQTAOPISNWSWYSLAGTNTKYQFSPSNYGPSTGYLYFPYKLVKADANNVGLQYTLN 364  
 QY 368 NGNVQVFEATSTSA-----NNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTSEG 422  
 Db 365 D-----ATKPAITFGSDQTMNGKTPYNDIVAVATLANLNFSGNKIEFSPVA--- 413  
 QY 423 NNNKVAPMIGNIYLSNENNADKI 446  
 Db 414 --EKVSPMIGNMYLSSSPNNMKI 435

RESULT 14

ID 005122 PRELIMINARY; PRT; 703 AA.  
 AC 005122;  
 DT 01-JUL-1997 (TREMBlrel. 04, Created)  
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Haemagglutinin.  
 GN PMGAL.9.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56;  
 RX MEDLINE=95010739; PubMed=7925999;  
 RA Markham P.F., Glew M.D., Sykes J.B., Bowden T.R., Pollocks T.D.,  
 RA Browning G.F., Whithear K.G., Walker I.D.,  
 RT "The organisation of the multigene family which encodes the major cell  
 RT surface protein, pmga, of Mycoplasma gallisepticum."  
 RL FEBS Lett. 352:347-352(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56;  
 RA Markham P.F.,  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U90714; AAB50154.1;  
 SQ SEQUENCE 703 AA; 75742 MW; 310B69BE9F73C8C5 CRC64;

Query Match 30.3%; Score 703.5; DB 2; Length 703;  
 Best Local Similarity 35.0%; Pred. No. 6e-26;  
 Matches 171; Conservative 77; Mismatches 125; Indels 115; Gaps 12;

QY 53 VGSFTYIRLEFGOMSTTKDAMP-----NNGOT-----OLEAA 84  
 Db 15 IGSFVMLAAAGTSTATTITPLNFTPNPBRKPDMPNPPSGGANGGNTNGMDTAOELASA 74  
 QY 85 RMELTDLINAKMTLASLDYAKIEASLSAYSEAEVNNINLATELQKAKTNLESAS 144  
 Db 75 KAALITLNRSEKVGIVDYAKIKADLTSAVTAKTSDSSTSLVQVKATSLQTAI 134  
 QY 145 NOANTDKTTPDNEHNVLEAKLTITEFOR-ATNLEGLSTAYNQIRNNIVDLYNKASS 203  
 Db 135 DKASDQKQFQDHDLMPESELTTLSQKNAATL--LNQPKGSAIINKINSIYAQSE 192  
 QY 204 LITKTLPLNGGTLDSNEITVNNINNTLS--TINOKTAPDALNSFIKVIQNNQ 261  
 Db 193 VVIRLDPVVS-GAITYAASITKNDIEIKAISENOJRKKNADAPFANYQPFKL---DXT 248  
 QY 262 SFVGTFTNANVOPSNYSFVAFSADVTVP-----NYKVARFTWNGDEPSSRIIANTN 313

Db 249 KIMGSTMNMKQPONYSFVGVSVGTGQWQSGOTTIPWNFQRIWMSGAPRADLASQTE 308  
 QY 314 S----- 314  
 Db 309 TPQAEPTPMASAPQCEVPAQOQGDSSPKQASFTQEVSPTPAAEVAQAQDTEQATSGCTP 368  
 QY 315 ITDVSWYSLAGTNTKYQFSPSNYGPSTGYLYFPYKLVKADANNVGLQYTLNNGVQV 374  
 Db 369 LTDVSWYSLAGTNTKYQFSPSNYGPSTGYLYFPYKLVKADANNVGLQYTLNNGVQV 426  
 QY 375 EPATSTANNTTANPTPAVDEIKVAKIVLSGLRFGQNTIELSVPTSEGNNKVAPMIGNI 434  
 Db 427 NFGSETNAN-----GPAASVDININAKVNLINLNGENTIERFSVP-----NNKVAPMIGNM 477  
 QY 435 YLSNENN 442  
 Db 478 YITSDVAN 485

RESULT 15

ID 09KH15 PRELIMINARY; PRT; 656 AA.  
 AC 09KH15;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE Adhesion PMGAL.2.  
 GN PMGAL.2.  
 OS Mycoplasma gallisepticum.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HS;  
 RA Shen Q.C., Bi D.R., Weng C.J.,  
 RT "Sequence analysis of the pmga multigene family of Mycoplasma  
 RT gallisepticum strain HS."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF25312; AAP91413.1;  
 SQ SEQUENCE 656 AA; 70875 MW; 58EA7B075FCC17B1 CRC64;

Query Match 30.2%; Score 703; DB 2; Length 656;  
 Best Local Similarity 40.7%; Pred. No. 5.9e-26;  
 Matches 166; Conservative 68; Mismatches 120; Indels 54; Gaps 13;

QY 71 DANPNNGOT-----OLEARMELTDLINAKMTLASLDYAKIEASLSAYSEAEVYN 123  
 Db 63 DINGGQNMWDSAAOELTARATLTSLASKNANVEYSDYAKIOWTLAAVTTAEGTS 122  
 QY 124 NNLAATLEOLKMAKTNESAINOANTDKTTPDNEHPMLVEAYKALKTTLLEGORATNLSGLS 183  
 Db 123 QNSATTEQVYNAATSAOATLNTANSNKQFQDQSHLSNLSYKULMATLAKKETAVMTLK 182  
 QY 184 STAYNOIRNNIVDLYNKASSLTITLPLNG-----GTLDSNEITVNNINNTLSGTIN 238  
 Db 183 DPKYSALIDQNGVSSKGBELVQHTLDPVSGIVPAANTI--TEBITKIEVISE--KTLQ 238  
 QY 239 EOKTNADALSN--SFT--KVIQNNQSFVGTFTNANVOPSNYSFVAFSADVTVP----- 289  
 Db 239 DQKNNAQOFANVQSFITLDTLLENVEDA-----KKMGQOPANYSFVGVSVDTGSGQET 292  
 QY 290 ---NYKVARFTWNGDEPSSRIIANTNS-----ITDVSWYSLAGTNTKYQFSPSNY 338  
 Db 293 TIPWNFAQRAIFTSNQOPTVYATTTGEGDOSTAKPLSDVSWYSLAGTNTKYQFSPSNY 352  
 QY 339 GPSTGYLYFPYKLVKADANNVGLQYTLN-NGNVQVFEATSTANNTTANPTPAVDEIK 397  
 Db 353 GPSTGYLYFPYKLVKAD--DVGQYKLNSEBELTPIIFBEGT---TTNCPATFVENIN 406  
 QY 398 VAKIVLSGLRFGQNTIELSVPTSEGNNKVAPMIGNIYLSNENNADK 445  
 Db 407 VAKVRLTGLAFGKNITIEFSVP-----MSKVAPMIGNMYITSSDYETNK 449

Mon Aug 25 09:49:57 2003

us-09-901-572a-3.rpt

Page 8

Search completed: August 14, 2003, 10:24:17  
Job time : 105 secs

---

THIS PAGE BLANK (USPTO)



C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D90011

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-2481 <KUR>

A:Cross-references: GB:BA000018; PID:gl3701961; PIDN:BA843253.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: fmcB(mrp)

Query Match 7.1%; Score 164.5; DB 2; Length 2481;  
Best Local Similarity 21.9%; Pred. No. 0.67;  
Matches 104; Conservative 66; Mismatches 173; Indels 131; Gaps 22;

```
QY 25 TONVTSREVSSVOLSEESTFYLCPPEVSGTVIRLEFGMSITKK--DANPNNGQTOL 81
Db 1150 TQDITAEIATA-----NADVDNAVTVQANSINBAANSQNDVDQAKTTGETSI 1196
QY 82 E-----AARMELDLINAKAMTLASIQDYAKIEASLSAYSEAEFTNNNLN---- 127
Db 1197 DQVTPYNNKATYANNEITAIINNKIQEIQATPDATDEKQADAA--EANTENGRANQAI 1254
QY 128 --ATLEQLKMAKTNLESAINQANTDKTFDNEHPNLYEAYKALKTTL---EQATNLEGL 182
Db 1255 AATTNAOVDEKAKNAEALIN-AVTPKVVKKQAKADEIDQLQATQTQTVNNDQNAINEKE 1313
QY 183 S-----STAYNQIINNLT-----VDLYNKASSLITKTLDPUNGTLIDSN-----E 222
Db 1314 AAIQOLATAVTDANKNITATATDNGVDPAKDAKNSIQSTOP--ATAVKSNAKNEVDQA 1370
QY 223 ITTVNRNINNTLSTINEOKTNADALNSFIKKVIQNNESFVGFTTNANVOPSNYSFYAF 282
Db 1371 VTTQOALDNTTGAATTEKNAKADL-----VLAKAKAYODIL--NAOTTN----- 1414
QY 283 SADVPVYKAYARRVWNGDEPSSRIILANTNSITDVSWIYSLA---GTNTKYQFSFENV 338
Db 1415 --DVTQI-----KQAAVADIGGITADTTIKQVAKDELATKANEEKAL--- 1454
QY 339 GPSTGYLYFPYKLVKADAN-----NVGLQYKLNNGNVQOVEPATISTANNTTA-NPT 390
Db 1455 -----IAQTADATTEKEKQANQOVDALTOGN-ONIEVAQSIDDVNTAKDNAI 1501
QY 391 PAVDEIKYAKIVLSGLRGONTIELSVPTGSGNMKVAPMIGNTYLSNENNAD 444
Db 1502 QALDPIDQASTDVKTAR-----AEILTEM---QNKITEIILNNNETTNEKGD 1546
```

Search completed: August 14, 2003, 10:22:28  
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2003, 01:32:14 ; Search time 5026 Seconds

(without alignments)  
10630.317 Million cell updates/sec

Title: US-09-901-572A-1

Perfect score: 1306

Sequence: 1 aaaaacacagatcgttcaat.....taaacgtttcaccaggtc 1306

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scb:\*  
12: gb\_ey:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_scb:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rtd:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_ey:\*  
39: em\_hggo\_hum:\*  
40: em\_hggo\_mus:\*  
41: em\_hggo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1280.4	98.0	2144	6 AR035279	AR035279 Sequence
2	1270	97.2	1387	6 AR035276	AR035276 Sequence
3	1264	96.8	1305	6 E09301	E09301 DNA encodin
4	1262.4	96.7	1387	6 E09301	E09301 DNA encodin
5	1179.4	90.3	302070	1 AE016968	AE016968 Mycoplasma
6	1086.8	83.2	2014	6 AR035278	AR035278 Sequence
7	1028	78.7	1152	6 AX65164	AX65164 Sequence
8	989.4	75.8	1015	6 E09896	E09896 Mycoplasma
9	986.2	75.5	1015	6 E09896	E09896 Mycoplasma
10	921.6	70.6	1082	6 AX65187	AX65187 Sequence
11	824.2	63.1	853	6 E02348	E02348 DNA sequenc
12	679.8	52.1	708	6 E02342	E02342 DNA sequenc
13	250.6	19.2	301042	1 AE016967	AE016967 Mycoplasma
14	249	19.1	4568	1 AE210770	AE210770 Mycoplasma
15	222.8	17.1	301903	1 AE016969	AE016969 Mycoplasma
16	217	16.6	7434	1 AE275312	AE275312 Mycoplasma
17	216.6	16.6	2369	6 AE035275	AE035275 Sequence
18	216.6	16.6	2369	6 E09895	E09895 Mycoplasma
19	215	16.5	3331	1 AE141940	AE141940 Mycoplasma
20	206.2	15.8	9720	1 MYCPMGAB	MYCPMGAB Mycoplasma
21	203.8	15.6	302070	1 AE016968	AE016968 Mycoplasma
22	199.8	15.3	3894	1 AY065985	AY065985 Mycoplasma
23	193.4	14.8	4715	1 AF032890	AF032890 Mycoplasma
24	191.8	14.7	8904	1 MGU90714	MGU90714 Mycoplasma
25	190.2	14.6	3897	1 AF053978	AF053978 Mycoplasma
26	179.2	13.3	2803	1 MYCHAHOM	MYCHAHOM Mycoplasma
27	173.2	13.3	2809	1 S55216	S55216 pMGAL.2-hem
28	136.4	10.4	702	6 E02344	E02344 DNA sequenc
29	92	7.0	170627	2 AC125567	AC125567 Ratrus no
30	89	6.8	145265	2 AL935272	AL935272 Danio rer
31	85	6.5	110000	3 AC116305_1	AC116305_1 Continuation (2 of
32	84.6	6.5	203138	2 BX530070	BX530070 Danio rer
33	82.4	6.3	94434	5 AL929250	AL929250 Zebrafish
34	82.2	6.3	155204	2 AC007926	AC007926 Trypanoso
35	80	6.1	154071	3 AC115598	AC115598 Dictyoste
36	79.2	6.1	153751	3 AC116551	AC116551 Dictyoste
37	79.2	6.1	225581	2 BX537105	BX537105 Danio rer
38	79	6.0	258920	2 BX510653	BX510653 Danio rer
39	78.8	6.0	178702	2 BX321884	BX321884 Danio rer
40	78.6	6.0	213544	2 BX510939	BX510939 Danio rer
41	78.2	6.0	594	6 E02345	E02345 DNA sequenc
42	77.4	5.9	335050	3 PFA929356	PFA929356 Plasmodiu
43	77.2	5.8	67970	3 PFMAL1P3	PFMAL1P3 Plasmodiu
44	76.4	5.8	2167	3 PRSTARA	PRSTARA P. relictow
45	75.8	5.8	2426	8 SDU49822	SDU49822 Saccharomyc

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AR035279	AR035279	Sequence 9 from patent US 5871742.	AR035279	AR035279.1	GI:5951947	Unknown.	Unknown.	1 (bases 1 to 2144)	Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.	Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine
AR035279	AR035279	2144 bp	DNA	linear	PAT 29-SEP-1999	Unknown.	Unknown.	1 (bases 1 to 2144)	Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.	Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine

JOURNAL Patent: US 5871742-A 9 16-FEB-1999;  
 FEATURES Location/Qualifiers

Source 1.2144  
 /organism="unknown"

BASE COUNT 811 a 341 c 332 g 660 t  
 ORIGIN

Query Match 98.0%; Score 1280.4; DB 6; Length 2144;  
 Best Local Similarity 98.8%; Pred. No. 7.9e-177;  
 Matches 1290; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 AAAACATCAGATGTTATCTGATATCTTCTTAAACACACAAATCTTACACA 60  
 Db 1 AAAACATCAGATGTTATCTGATATCTTCTTAAACACACAAATCTTACACA 60  
 QY 61 AATCTTAATAATAAGCCGTTAATTAATAATAATAATAATAATAATAATAATA 120  
 Db 61 AATCTTAATAATAAGCCGTTAATTAATAATAATAATAATAATAATAATAATA 120  
 QY 121 AACCAAAATCTCTGATATAACGCTTATTTATTTTATTTTATTTTATTTTATTT 180  
 Db 121 AACCAAAATCTCTGATATAACGCTTATTTATTTTATTTTATTTTATTTTATTT 180  
 QY 181 ATAATATATCTTAATATCTATGATAAAGAAAGATCATCTTAAGACTATAGTTG 240  
 Db 181 ATAATATATCTTAATATCTATGATAAAGAAAGATCATCTTAAGACTATAGTTG 240  
 QY 241 TTAGGTACACATCCTTTCTTAGCATTTGGATTTCTAGCTGATGCTATTAATAAAA 300  
 Db 241 TTAGGTACACATCCTTTCTTAGCATTTGGATTTCTAGCTGATGCTATTAATAAAA 300  
 QY 301 GATGAAACCCAAATTAATGCGCAATTTAGAAAGCGCGGAAATGAGTTACAAAT 360  
 Db 301 GATGAAACCCAAATTAATGCGCAATTTAGAAAGCGCGGAAATGAGTTACAAAT 360  
 QY 361 CTAAATCAATGCTAAAGCGATGACATTTAGCTTCACTACAAAGCTATAGCCAAATTAAGCT 420  
 Db 361 CTAAATCAATGCTAAAGCGATGACATTTAGCTTCACTACAAAGCTATAGCCAAATTAAGCT 420  
 QY 421 AGTTATCATCTGCTTATTAAGTGAAGCTGAACAGTTAACATTAACCTTATGCAACACTA 480  
 Db 421 AGTTATCATCTGCTTATTAAGTGAAGCTGAACAGTTAACATTAACCTTATGCAACACTA 480  
 QY 481 GAACAACCTAAATGCTAAATCTAATTTAGAAATGAGCATTAACCAAGCTAATAGGAT 540  
 Db 481 GAACAACCTAAATGCTAAATCTAATTTAGAAATGAGCATTAACCAAGCTAATAGGAT 540  
 QY 541 AAAAGACTTTGATTAATGAACACCAATTTAGTTGAAGCATTAAGCACTAAACAC 600  
 Db 541 AAAAGACTTTGATTAATGAACACCAATTTAGTTGAAGCATTAAGCACTAAACAC 600  
 QY 601 ACTTTAGAACACAGTGTACTAATCTTGAAGGTTTGTCTCAACGCTTATTAATCAAT 660  
 Db 601 ACTTTAGAACACAGTGTACTAATCTTGAAGGTTTGTCTCAACGCTTATTAATCAAT 660  
 QY 661 CGCAATTAATTTAGTGTATTAACAAATAGCTAGTAACTTAATACTTAACACTAGAT 720  
 Db 661 CGCAATTAATTTAGTGTATTAACAAATAGCTAGTAACTTAATACTTAACACTAGAT 720  
 QY 721 CCACTAAATGAGGGAACGCTTTAGATTTCTATGAGATTAATAGTTATGGAATAT 780  
 Db 721 CCACTAAATGAGGGAACGCTTTAGATTTCTATGAGATTAATAGTTATGGAATAT 780  
 QY 781 AATAATAGCTTAATCACTATTAATGAACAAAGACTAATGCTAGCATTAATCAATAGT 840  
 Db 781 AATAATAGCTTAATCACTATTAATGAACAAAGACTAATGCTAGCATTAATCAATAGT 840  
 QY 841 TTTATTAATAAAGTATTAATAATGAACAAAGTTTGTAGGAGCTTTTCAAAAGCT 900  
 Db 841 TTTATTAATAAAGTATTAATAATGAACAAAGTTTGTAGGAGCTTTTCAAAAGCT 900  
 QY 901 AATGTTCAACCTTCAACACTAGTTTGTGCTTTTATGCTGATGAACACCCGTCAT 960  
 Db 901 AATGTTCAACCTTCAACACTAGTTTGTGCTTTTATGCTGATGAACACCCGTCAT 960

Db 901 AATGTTCAACCTTCAACACTAGTTTGTGCTTTTATGCTGATGAACACCCGTCAT 960  
 QY 961 TTTAATATGCAAGAGACCGTTTGGATGATGAACTTCAAGTAAATTTTTCGA 1020  
 Db 961 TTTAATATGCAAGAGACCGTTTGGATGATGAACTTCAAGTAAATTTTTCGA 1020  
 QY 1021 AACAGAAATGATACAGATGTTTCTGATTAATTAAGTTATAGCTGAACAAACGAG 1080  
 Db 1021 AACAGAAATGATACAGATGTTTCTGATTAATTAAGTTATAGCTGAACAAACGAG 1080  
 QY 1081 TACCAATTTAGTTTATGCAATATGTCATCAACTGCTTAATTAATTTCCCTTAAG 1140  
 Db 1081 TACCAATTTAGTTTATGCAATATGTCATCAACTGCTTAATTAATTTCCCTTAAG 1140  
 QY 1141 TTGGTTAAGACGCTGATGCTTAATTAAGTTTATCAATCAATTAATTAATGAAT 1200  
 Db 1141 TTGGTTAAGACGCTGATGCTTAATTAAGTTTATCAATCAATTAATTAATGAAT 1200  
 QY 1201 GTTCAACAAGTTGATGCTTCCACTTCAACTAGTCAAAATAATTAATCACTAATCACT 1260  
 Db 1201 GTTCAACAAGTTGATGCTTCCACTTCAACTAGTCAAAATAATTAATCACTAATCACT 1260  
 QY 1261 CCAGCAGTTGATGAGATTTAAAGTTGCTAAATTCGTTTATCAGGTT 1306  
 Db 1261 CCAGCAGTTGATGAGATTTAAAGTTGCTAAATTCGTTTATCAGGTT 1306

RESULT 2  
 AR035276 1387 bp DNA linear PAT 29-SEP-1999  
 LOCUS AR035276  
 DEFINITION Sequence 3 from patent US 5871742.  
 ACCESSION AR035276  
 VERSION AR035276.1 GI:5951944  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1387)  
 AUTHORS Saitoh,S., Ohkawa,S., Saeki,S., Ohsawa,I., Funato,H., Iritani,Y., Aoyama,S. and Takahashi,K.  
 TITLE Recombinant Avipox virus encoding polypeptide of mycoplasma gallisepticum, and utilized a live vaccine  
 JOURNAL Patent: US 5871742-A 3 16-FEB-1999;  
 FEATURES Location/Qualifiers  
 source 1.1387  
 BASE COUNT 531 a 220 c 207 g 429 t

Query Match 97.2%; Score 1270; DB 6; Length 1387;  
 Best Local Similarity 98.8%; Pred. No. 3e-175;  
 Matches 1290; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1 AAAACATCAGATGTTATCTGATATCTTCTTAAACACACAAATCTTACACA 60  
 Db 1 AAAACATCAGATGTTATCTGATATCTTCTTAAACACACAAATCTTACACA 60  
 QY 61 AATCTTAATAATAAGCCGTTAATTAATAATAATAATAATAATAATAATAATA 120  
 Db 61 AATCTTAATAATAAGCCGTTAATTAATAATAATAATAATAATAATAATAATA 120  
 QY 121 AACCAAAATCTCTGATATAACGCTTATTTATTTTATTTTATTTTATTTTATTT 180  
 Db 121 AACCAAAATCTCTGATATAACGCTTATTTATTTTATTTTATTTTATTTTATTT 180  
 QY 181 ATAATATATCTTAATATCTATGATAAAGAAAGATCATCTTAAGACTATAGTTG 240  
 Db 181 ATAATATATCTTAATATCTATGATAAAGAAAGATCATCTTAAGACTATAGTTG 240  
 QY 241 TTAGGTACACATCCTTTCTTAGCATTTGGATTTCTAGCTGATGCTATTAATAAAA 300  
 Db 241 TTAGGTACACATCCTTTCTTAGCATTTGGATTTCTAGCTGATGCTATTAATAAAA 300

```

QY 301 GATGCAAAACCCAAATATGCGCCAAACCCAAATTAGACGCCGGAATGAGTTAAACAGAT 360
DB 301 GACGCAAAACCCAAATATGCGCCAAACCCAAATTAGACGCCGGAATGAGTTAAACAGAT 360
QY 361 CTAAATCAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 CTAAATCAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 AGTTATCATCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 AGTTATCATCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GAAACAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GAAACAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 AAAAGGACTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 AAAAGGACTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 ACTTTAGAACAAACGCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ACTTTAGAACAAACGCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 CGCAATTAATTTAGTGAATCTATACAAATTAAGCTAGTGAATTAATTAATTAATTAATTAAT 720
DB 661 CGCAATTAATTTAGTGAATCTATACAAATTAAGCTAGTGAATTAATTAATTAATTAATTAAT 720
QY 721 CCATTAATTTGGGGGAAACGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 CCATTAATTTGGGGGAAACGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 AATATAGCTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
DB 781 AATATAGCTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
QY 841 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
DB 841 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 901 AATGTTCAACCTTCAAACTACAGCTTTTGTGCTTTTGTGCTTTTGTGCTTTTGTGCTTTTGTGCT 960
DB 901 AATGTTCAACCTTCAAACTACAGCTTTTGTGCTTTTGTGCTTTTGTGCTTTTGTGCTTTTGTGCT 960
QY 961 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
DB 961 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
QY 1021 AACACGAATAGTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 AACACGAATAGTATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
DB 1081 TACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
QY 1141 TTGGTTAAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 TTGGTTAAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 GTTCAACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GTTCAACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CCAGCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
DB 1261 CCAGCAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306

```

RESULT 3  
 LOCUS 117387 1305 bp DNA linear PAT 06-OCT-1996  
 DEFINITION Sequence 1 from patent US 5489430.

```

ACCESSION 117387
VERSION 117387.1 GI:1597742
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1305)
AUTHORS Saito,S., Ohkawa,S., Fujisawa,A., Iritani,Y. and Aoyama,S.
TITLE Polytomy mycoplasma antigen, gene thereof and recombinant vectors
JOURNAL containing the gene as well as vaccines utilizing the same
FEATURES
    Patent: US 5489430-A 1 06-FEB-1996;
    Location/Qualifiers
        1..1305
BASE COUNT 499 a 208 c 184 g 408 t 6 others
ORIGIN
Query Match 96.8%; Score 1264; DB 6; Length 1305;
Best Local Similarity 98.3%; Pred. No. 2.2e-174;
Matches 1284; Conservative 0; Mismatches 21; Indels 1; Gaps 1;
QY 1 AAAAATCATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
DB 1 AAAAATCATGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 61 AATCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
DB 61 AATCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 120
QY 121 AACCAAAATCTAGTAAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 180
DB 121 AACCAAAATCTAGTAAATTAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 180
QY 181 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
DB 181 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
QY 241 TTAGTCAACATGCTTTCTAGCATTTGGGATTTCTAGCATTTGGGATTTCTAGCATTTGGGATTT 300
DB 241 TTAGTCAACATGCTTTCTAGCATTTGGGATTTCTAGCATTTGGGATTTCTAGCATTTGGGATTT 300
QY 301 GATGCAAAACCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
DB 301 GATGCAAAACCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
QY 361 CTAAATCAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 CTAAATCAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 AGTTATCATCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 AGTTATCATCTGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 GAAACAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GAAACAATGCTAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 AAAAGGACTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 AAAAGGACTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 ACTTTAGAACAAACGCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 ACTTTAGAACAAACGCTGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 CGCAATTAATTTAGTGAATCTATACAAATTAAGCTAGTGAATTAATTAATTAATTAATTAAT 720
DB 661 CGCAATTAATTTAGTGAATCTATACAAATTAAGCTAGTGAATTAATTAATTAATTAATTAAT 720
QY 721 CCATTAATTTGGGGGAAACGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 CCATTAATTTGGGGGAAACGCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 780

```

QY 781 AATAATACGTTATCACTATTAATGACAAAGACTAGTGCATTAATCTATAGT 840  
DB 781 AATAATACGTTATCACTATTAATGACAAAGACTAGTGCATTAATCTATAGT 840  
QY 841 TTTATTAATAAAGTGAATCAAAATATGACAAAGTTTGTGAGGACCTTTTACAAACGCT 900  
DB 841 TTTATTAATAAAGTGAATCAAAATATGACAAAGTTTGTGAGGACCTTTTACAAACGCT 900  
QY 901 AATGTTCAACCTTCAACACTAGCTTTGTTGCTTTTGTAGGCTGATGTAACCCGCTCAT 960  
DB 901 AATGTTCAACCTTCAACACTAGCTTTGTTGCTTTTGTAGGCTGATGTAACCCGCTCAT 960  
QY 961 TATTAATATGACAAAGACCGCTTNNNATGATGTAACCTTCAAGTGAATTTCTTGA 1020  
DB 961 TATTAATATGACAAAGACCGCTTNNNATGATGTAACCTTCAAGTGAATTTCTTGA 1020  
QY 1021 AACAGAAATGATCAAGATGTTTCTTGATTTATGTTAGCTGGAACAAACAGAG 1080  
DB 1021 AACAGAAATGATCAAGATGTTTCTTGATTTATGTTAGCTGGAACAAACAGAG 1080  
QY 1081 TACCAATTTAGTTTACGAACTAGTGCATGCACTGTTATTTATTTCCCTATAG 1140  
DB 1081 TACCAATTTAGTTTACGAACTAGTGCATGCACTGTTATTTATTTCCCTATAG 1140  
QY 1141 TTGTTTAAAGCAGCTGATGCTAATACGTTGATTAACAATAATTAATGGAAT 1200  
DB 1141 TTGTTTAAAGCAGCTGATGCTAATACGTTGATTAACAATAATTAATGGAAT 1200  
QY 1201 GTTCAACAAGTTGAGTTTCCCACTTCACTAGTGAATAATACGCTATGCAACT 1260  
DB 1201 GTTCAACAAGTTGAGTTTCCCACTTCACTAGTGAATAATACGCTATGCAACT 1260  
QY 1261 CCAGCAGTTGATGATTAAGTTGCTAAATCGTTTATCAGTT 1306  
DB 1261 CCAGCAGTTGATGATTAAGTTGCTAAATCGTTTATCAGTT 1306

RESULT 4  
E09301 1387 bp DNA linear PAT 29-SEP-1997  
LOCUS E09301  
DEFINITION DNA encoding Mycoplasma TM-16.  
ACCESSION E09301.1 GI:22025928  
VERSION JP 1995133295-A/2.  
KEYWORDS Mycoplasma gallisepticum  
SOURCE Mycoplasma gallisepticum  
ORGANISM Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
REFERENCE 1 (bases 1 to 1387)  
AUTHORS Mori,H., Saito,S., Okawa,S., Funato,H., Iritani,K., Aoyama,S. and Takahashi,K.  
TITLE NEW ANTIGEN PROTEIN, ITS GENE, RECOMBINANT BACULOVIRUS AND ITS USE  
JOURNAL Patent: JP 1995133295-A 2 23-MAY-1995.  
COMMENT NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
OS Mycoplasma gallisepticum  
PN JP 1995133295-A/2  
PD 23-MAY-1995  
PF 27-AUG-1993 JP 1993213102  
PI MORI HAJIME, SAITO SHUJI, OKAWA SEISUKO, FUNATO HIRONO, PI  
IRITANI KOICHI,  
PI AOKIYAMA SHIGEMI, TAKAHASHI KIYOTO  
PC C07K14/30,A61K39/00,A61K39/00,C12N7/01,C12N15/31,C12P21/02, PC  
(C12P21/02,  
PC C12R1:92);  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FT source 1..1387  
FT 202..1308 /organism='Mycoplasma gallisepticum' FT CDS  
FT Location/Qualifiers  
FT 1..1387

FEATURES  
source 1..1387

BASE COUNT 531 a 221 c 202 g 427 t 6 others  
ORIGIN  
Query Match 96.7%; Score 1262.4; DB 6; Length 1387;  
Best Local Similarity 98.2%; Pred. No. 3.7e-174;  
Matches 1283; Conservative 0; Mismatches 22; Indels 1; Gaps 1;  
/organism="Mycoplasma gallisepticum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2096"  
1 AAAAATCATGATGTTATCTGATCTTGTCTTAAAAAACAACAAATCTCTACAA 60  
1 AAAAATCATGATGTTATCTGATCTTGTCTTAAAAAACAACAAATCTCTACAA 60  
61 AATCTTAATTAATGAGCGTTAATTAATTAATTAATTAATTAATTAATTAATTA 120  
61 AATCTTAATTAATGAGCGTTAATTAATTAATTAATTAATTAATTAATTAATTA 120  
121 AACCAAAATCTCTAGTAAATGAGCTTATTTATTTATTTATTTATTTATTTAT 180  
121 AACCAAAATCTCTAGTAAATGAGCTTATTTATTTATTTATTTATTTATTTAT 180  
181 AATAATATCTTAATATCTTATGATTAAGAAAGAAATCACTTAAGACTATTA 240  
181 AATAATATCTTAATATCTTATGATTAAGAAAGAAATCACTTAAGACTATTA 240  
241 TTAGTCAACATCTTCTTATGAGATGGGATTTCTAGCTGATGCTATTAATTA 300  
241 TTAGTCAACATCTTCTTATGAGATGGGATTTCTAGCTGATGCTATTAATTA 300  
301 GATGCAAAACCAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 360  
301 GATGCAAAACCAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 360  
361 GAGCAAAACCAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 420  
361 GAGCAAAACCAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 420  
421 AGTTATCATCTGTTATGAGAGCTGAAACCTTAACAATCACTTAATGAAACATA 480  
421 AGTTATCATCTGTTATGAGAGCTGAAACCTTAACAATCACTTAATGAAACATA 480  
481 GAAACAAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 540  
481 GAAACAAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 540  
541 GAAACAAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 600  
541 GAAACAAATTAATGAGCAACCAATTAAGAGAGCGCAATGAGTTAACA 600  
601 ACTTGAACAAGAGTCTACCTTGAAGGTTGTCATCACTGCTTAATCAAT 660  
601 ACTTGAACAAGAGTCTACCTTGAAGGTTGTCATCACTGCTTAATCAAT 660  
661 ACTTGAACAAGAGTCTACCTTGAAGGTTGTCATCACTGCTTAATCAAT 720  
661 ACTTGAACAAGAGTCTACCTTGAAGGTTGTCATCACTGCTTAATCAAT 720  
721 CCATTAATGAGGAGAGCGCTTTGATTTCTAATGAGTACTCAAGTAATCGGAAT 780  
721 CCATTAATGAGGAGAGCGCTTTGATTTCTAATGAGTACTCAAGTAATCGGAAT 780  
781 AATAATGTTATGATCACTTAATGACAAAGCTAATGCTGATTAATCTAATAGT 840  
781 AATAATGTTATGATCACTTAATGACAAAGCTAATGCTGATTAATCTAATAGT 840  
841 TTTATTAATAAAGTGAATCAAAATATGACAAAGTTTGTGAGGACCTTTTACAA 900  
841 TTTATTAATAAAGTGAATCAAAATATGACAAAGTTTGTGAGGACCTTTTACAA 900  
901 AATGTTCAACCTTCAACACTAGCTTTGTTGCTTTTGTAGGCTGATGTAACCCG 960  
901 AATGTTCAACCTTCAACACTAGCTTTGTTGCTTTTGTAGGCTGATGTAACCCG 960

QY 961 TATAAATATGCAAGAGACCGTTTGATGATGATGATGAACCTTCAAGTAGAATTCCTTCA 1020  
 Db 961 TATAAATATGCAAGAGACCGTTTNNATATGATGATGAACCTTCAAGTAGAATTCCTTCA 1020  
 QY 1021 AACACGATAGTATACAGATGTTTCTTGATTTTACTGTTAGCTGGAACAAACGAG 1080  
 Db 1021 AACACGATAGTATACAGATGTTTCTNNATTTTACTGTTAGCTGGAACAAACGAG 1080  
 QY 1081 TACCAATTAGTTTAGCAACTATGATGATCACTGATTTATTTATTCCTTATAG 1140  
 Db 1081 TACCAATTAGTTTATGACACTATGATGATCACTGATTTATTTATTCCTTATAG 1140  
 QY 1141 TTGGTTAAGACAGCTGATGCTAATACGTTGATTTACATCAATTAATTAATGGAAT 1200  
 Db 1141 TTGGTTAAGACAGCTGATGCTAATACGTTGATTTACATCAATTAATTAATGGAAT 1200  
 QY 1201 GTTCAACAGTTGAGTTGCCACTTCAACTGATGCAATTAATATACAGCTAATCAACT 1260  
 Db 1201 GTTCAACAGTTGAGTTGCCACTTCAACTGATGCAATTAATATACAGCTAATCAACT 1260  
 QY 1261 CCAGAGTTGATGATTAAGTTGCTAATTCGTTTATACAGTT 1306  
 Db 1261 -CAGCAGTTGATGATTAAGTTGCTAATTCGTTTATACAGTT 1305  
 RESULT 5  
 AEO16968 302070 bp DNA linear BCT 09-JUN-2003  
 LOCUS DEFINITION Mycoplasma gallisepticum strain R section 2 of 4 of the complete genome.  
 ACCESSION AEO16968 AEO15450  
 VERSION AEO16968.1 GI:31541264  
 KEYWORDS  
 SOURCE Mycoplasma gallisepticum R  
 ORGANISM Mycoplasma gallisepticum R  
 Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 REFERENCE 1 (bases 1 to 302070)  
 AUTHORS Geary, S.J., Papazisi, L., Kutish, G., Gorton, T.S., Mahairas, G., Swartzell, S., Madan, A., Nguyen, D.K., Matham, P., Browning, G., Kamal, M. and Liao, X.  
 TITLE The complete genome sequence of the avian pathogen Mycoplasma gallisepticum strain R  
 JOURNAL Microbiology (2003) In press  
 REFERENCE 2 (bases 1 to 302070)  
 AUTHORS Geary, S.J., Papazisi, L., Kutish, G., Mahairas, G., Swartzell, S., Madan, A., Nguyen, D.K., Gorton, T.S., Matham, P., Browning, G., Mustata, K. and Liao, X.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-OCT-2002) Department of Pathobiology and Veterinary Sciences, and Center of Excellence for Vaccine Research, The University of Connecticut, 61 North Eagleville Road U-89, Storrs, CT 06269-3089, USA  
 FEATURES  
 source 1. 302070  
 /organism="Mycoplasma gallisepticum R"  
 /mol\_type="Genomic DNA"  
 /strain="R"  
 /db\_xref="taxon:233150"  
 79. 3588  
 /locus\_tag="MGA\_1011"  
 79. 3588  
 /locus\_tag="MGA\_1011"  
 /locus\_tag="MGA\_1011"  
 /note="MGA\_1011; MGR\_224"  
 /codon\_start=1  
 /transl\_table=4  
 /product="conserved hypothetical"  
 /protein\_id="AAP5656.1"  
 /db\_xref="GI:31541265"  
 /translation="MMNNKKNNPKKPLVKNKGLDLDPISEIIVLNDPHLTKPKK  
 QQPLLEKRAKKVQPVVYKGEKKAKAKAVITDPRLTKGSRLLITIGFLIG  
 GITPULSFALINOIDKOTDOSAPITRISDPOFTHDNLSPFVNSYOL  
 NKOSINGIDISNORTLHDLPIASFMINKLVRLNOLHFNYSMLYSDIAYLRKTY

DENKAFIPKTTDLOITKSSDNNPLHVLTFANEFNQTDSPTQFVFRFNORYEKO  
 IOPKGINILKLSASANDIDPLKDPYISIDVDLAPGYINAVYKOLYHNYLL  
 NMLNSNLEAMNDENYRSENFENNNPNSMILTKTIPGDPYQAOANDVGENKYL  
 RDLTPBOISNIRKTFEENPHSOEEMNLRDVKKMLVNNNKKXLYVSDAAAF  
 AKLPSVFIKILNRYARSKAILKULIDALSSNFTYLNKAVIEVLONDFKULIL  
 TYFDFVASYNLIDSLTKDVNSAFDFQRYELMPLROLISNFPYKIFPEILERIL  
 KYEDDADNLPPIPEFNTRYLINTIPESRLLNDYDOLNINSDVPMATANTNNGSS  
 MSRETRLENFLFIONRSTLVNLFQSYNPEDDLTLFISNNFPLPOLKLFGINVAT  
 LDLEFKTALDILFLKNPEILSRDLKAKIINLFPMWIIKSLDQTNINKIKELGYKRE  
 DNVLNOSPSTLVKXKLDLGPKLDINGLOIYGVLSAFERLPYKIVDFPKYVDDK  
 RYERLKEKGYQOALNEGSSULNAGGIFSRKMGADPAKNFSSVFDDIWKRLIKGVS  
 DNNETKIDLANEIPSTGNNDPNKVIATTINGAGYFKVSGENVLDLPEELIPSNINQ  
 QIVYQLVNDOEKPQIYQNLKLFERYSRQGFPSDSLKSKIYQISISQWQITON  
 YSLNRKYLITFTRILILPGDGSKNHLIQENNPYENVALFTYNAOSILNOYDRANED  
 KRRLREOILMITDRBARVKGIDRPKRETSILPOSRLMDSLIOIGVNSLYIHTN  
 MSPTISERSISLNTTEVRKGLISKQVSNLAYETPLSASVQVLVYPIILDSNPN  
 PVFKKEISHLVSLFPGAGE"  
 3664. 5019  
 /locus\_tag="MGA\_1014"  
 3664. 5019  
 /locus\_tag="MGA\_1014"  
 /locus\_tag="MGA\_1014; MGR\_225"  
 /note="MGA\_1014; MGR\_225"  
 /codon\_start=1  
 /transl\_table=4  
 /product="unique hypothetical"  
 /protein\_id="AAP5657.1"  
 /db\_xref="GI:31541266"  
 /translation="MTPEYKKEIPDYDEGVDSIFHNDERTDPEDDQDHRGFSRTR  
 RIKKNVYDQHDLTLDVSRKKNRFEKPLVRSTITKTFIAMLWLILFIVLIVF  
 VLNKFSFQOSESITYSVNPLPTJALAAVSFII SVILGFMVNHKQIKYSITERSLDF  
 YSIFPEKLYVESYFLGMIPLGLIVISVSHLQNLDDSDVSDNNLMOESISDLIDM  
 NYGNSYRLESKKVYLDQELQRLNRYKOKLAKNKLAVVSAALNPFKSKOLLYN  
 VNSGFSASLIVQVKSATITOLREBRKLRBRKRLKALISNDRALIDEQPLIKETRO  
 LPDGBESLFINLIDPSRKNREGNLENFDRIDMISENLNDRALIVEIGIKKETRO  
 NLSRAPKVVKNKQPTFKVKNVNNKKAALVQROESDLTKOSNLIFVARIWVQ  
 L"  
 5196. 5269  
 /product="tRNA-Trp"  
 /note="codon recognized: UGC; tRNA genes predicted using  
 PASTRNA software of the Institut Pasteur, France;  
 functional prediction of the tRNA is performed according  
 to Ushida C. and Muto"  
 6561. 7661  
 /gene="hacA"  
 /locus\_tag="MGA\_1017"  
 6561. 7661  
 /gene="hacA"  
 /locus\_tag="MGA\_1017"  
 /note="MGA\_1017; high affinity transport protein A (HacA)  
 [Mycoplasma gallisepticum]; MGR\_226"  
 /codon\_start=1  
 /transl\_table=4  
 /product="HacA"  
 /protein\_id="AAP5658.1"  
 /db\_xref="GI:31541267"  
 /translation="MSKTIKKLPLISSPILSSAFVSSCAKANDANNQGTLOLVKDK  
 IDHDAVALINDYIKNLSQYVSKLNKVSITSNTERGQSSIDPKSNLIDLAVSS  
 AKTNDYDITTKIOTLRSFRFDTPPTFAGTLDQAKIKSILFNEIPRNMWDER  
 MMDGNKYQFFGPEDDLISFRKMILIGQODTLAKIKEMAQKNMVFSSYGGGR  
 TSSNGRYTYPNLNKKHNDPNLTIISQVVRGEIGCTNRPVHIVPDDMSFATQNK  
 NKVNPYTFNPNNDNTKVEILSLDPLALNIGVFNNKRLTMDVINAVESIYELAKNNO  
 DTVGPTVGVNGYRIKDVNKEVYQFIKANIN"  
 7661. 8407  
 /gene="hacB"  
 /locus\_tag="MGA\_1018"  
 7661. 8407  
 /gene="hacB"  
 /locus\_tag="MGA\_1018"  
 /note="MGA\_1018; high affinity transport protein B (HacB)  
 [Mycoplasma gallisepticum]; MGR\_227"  
 /codon\_start=1  
 /transl\_table=4  
 /product="HacB"

Query Match	90.3%	Score 1179.4	DB 1	Length 302070
Best Local Similarity	95.3%	Pred. No. 7.2e-163		
Matches 1249	Conservative	0	Mismatches 56	Indels 6
			Gaps 3	
	<pre> /product='unique hypothetical' /protein_id='AAP56573.1' /db_xref='GI:31541272' /translation='MULTISUBSTRATE HYDROLASE' PIRERLALSDVNDNNLNLFEILENNISILDPNTQSNVFLAGNYSVSEIT NEETIDQKQFQVSPLEKTREREESIDISVIDLYQNDLIEGRDLSNLR LLEENKDALVTKIKLNNLEFRDVRGDRIDADYQVEKKAISWQPNVPLSND QINDLTLINVDVAVTLFKSRVVEDYTSITNGSKSLVIENTNOITLYPNDIKLALDT </pre>			
QY	1	AAAAACATCGATGTTGATATCTGATATCTCTTGGCTTAAAAAACAACAAATCTTCTAACAA	60	
DB	189913	AGAACATCGATGTTGATATCTGATATCTCTTGGCTTAAAAAACAACAAATCTTCTAACAA	189971	
QY	61	AATCTTAATTAATTAAGCCGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	120	
DB	189972	AATCTTAATTAATTAAGCCGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	190031	
QY	121	AACCAAAATTCCTAGATTAATTAAGCCGTTATATTTATTTTATTTAGTCATCTTTTAA	179	
DB	190032	AACCAAAATTCCTAGATTAATTAAGCCGTTATATTTATTTTATTTAGTCATCTTTTAA	190091	
QY	180	TA---TAAATATATCTTAATTAATTTCTATGTAATTAAGAAAGATCATCTTAAAGACTATTA	235	
DB	190092	TAAATATATCTTAATTAATTTCTATGTAATTAAGAAAGATCATCTTAAAGACTATTA	190151	
QY	236	GTTTGTATAGTGAACAACATCCCTTTCTAGATGAGGAGATTTCTAGCTGATGCTATTA	295	
DB	190152	GTTTGTATAGTGAACAACATCCCTTTCTAGATGAGGAGATTTCTAGCTGATGCTATTA	190211	
QY	296	AAAAAGATGCAAAACCCAAATTAATGAGCCAAACCAATTAAGAGCGCGAATGAGTTAA	355	
DB	190212	AAAAAGATGCAAAACCCAAATTAATGAGCCAAACCAATTAAGAGCGCGAATGAGTTAA	190271	
QY	356	CAGATCTATCAATGCTTAAGCCATGATCTTACCTTCACTAACAAGCTATGCGAAGATTG	415	
DB	190272	CAGATCTATCAATGCTTAAGCCATGATCTTACCTTCACTAACAAGCTATGCGAAGATTG	190331	
QY	416	AAGCTAGTTTCAATCTGCTATAGTAAAGCTGAAACAGTTTAACTTAATCTTAAATGCAA	475	
DB	190332	AAGCTAGTTTCAATCTGCTATAGTAAAGCTGAAACAGTTTAACTTAATCTTAAATGCAA	190391	
QY	476	CATTAGAACAACTAAAAATGCGTAAATCTAATTTAGATAGCCATCAACCAAGCTAAAT	535	
DB	190392	CATTAGAACAACTAAAAATGCGTAAATCTAATTTAGATAGCCATCAACCAAGCTAAAT	190451	
QY	536	CGGATTAACCACTTTGATATTAATGAACCCCAATTTAGTTGAAGCTAACCAAGCACTAA	595	
DB	190452	CGGATTAACCACTTTGATATTAATGAACCCCAATTTAGTTGAAGCTAACCAAGCACTAA	190511	
QY	596	AAACCACTTTAGAACAAAGTCTCTAATCCTTGAAGGTTTGCTCACTGCTTATATC	655	
DB	190512	AAACCACTTTAGAACAAAGTCTCTAATCCTTGAAGGTTTGCTCACTGCTTATATC	190571	
QY	656	AAATTCGCAATTAATTTAGATCTATTAATGAACAAAGCTAGTATTAATTAATTAATTAAC	715	
DB	190572	AAATTCGCAATTAATTTAGATCTATTAATGAACAAAGCTAGTATTAATTAATTAATTAAC	190631	
QY	716	TAGATCACTAATGAGGGAACGCTTTAGATTTCTAATGAGATTCTACAGTTATCGGA	775	
DB	190632	TAGATCACTAATGAGGGAACGCTTTAGATTTCTAATGAGATTCTACAGTTATCGGA	190691	
QY	776	ATATTAATTAATCGTTATCACTATTAATGAACAAAGCTAATGCTAGATGCACTAATCTA	835	
DB	190692	ATATTAATTAATCGTTATCACTATTAATGAACAAAGCTAATGCTAGATGCACTAATCTA	190751	
QY	836	ATATGTTTATTAATAAAGATTAATCAAAATTAATGAACAAAGTTTGTGAGGACTTTTACAA	895	
DB	190752	ATATGTTTATTAATAAAGATTAATCAAAATTAATGAACAAAGTTTGTGAGGACTTTTACAA	190811	



RESULT	7			
LOCUS	AX665164			
DEFINITION	Sequence 1 from Patient EPI275716.	1152 bp	DNA	linear
ACCESSION	AX665164			PAT 26-MAR-2003
VERSION	AX665164.1	GI:29290294		
KEYWORDS				
SOURCE	Mycoplasma gallisepticum			
ORGANISM	Mycoplasma gallisepticum			



QY 121 AACCAAAATCTCTAGTATATAACCGTTATTTATTTATTTAGTCATCTTTAGAT 180  
 DB 121 AACCAAAATCTCTAGTATATAACCGTTATTTATTTATTTAGTCATCTTTAGAT 180  
 QY 181 ATTAATATATCTTATATTTCTTGAATAAGAAAGATCATCTTAAAGACTTTAGTTG 240  
 DB 181 ATTAATATATCTTATATTTCTTGAATAAGAAAGATCATCTTAAAGACTTTAGTTG 240  
 QY 241 TTAGGTACAAACCTCTTCTTGAATTTAGGATTTCTAGCTGATCTTATTAATAAAA 300  
 DB 241 TTAGGTACAAACCTCTTCTTGAATTTAGGATTTCTAGCTGATCTTATTAATAAAA 300  
 QY 301 GATGCAAAACCAAAATATGCGCAAAACCAATTTAGAGAGCGGATGAGATTAACAGAT 360  
 DB 301 GATGCAAAACCAAAATATGCGCAAAACCAATTTAGAGAGCGGATGAGATTAACAGAT 360  
 QY 361 CTATATCATGCTTAAAGCGATGACATTAGCTTACATCAAGACTATGCCAAGTTGAAGCT 420  
 DB 361 CTATATCATGCTTAAAGCGATGACATTAGCTTACATCAAGACTATGCCAAGTTGAAGCT 420  
 QY 421 AGTTATCATGCTTATAGTGAAGCTGAAACAGTTAACATTAACCTTAAATGCAACATTA 480  
 DB 421 AGTTATCATGCTTATAGTGAAGCTGAAACAGTTAACATTAACCTTAAATGCAACATTA 480  
 QY 481 GAACAACCTAAATGCTTAAATCTAATTTAGATGAGCATCAACCAAGCTAATACGAT 540  
 DB 481 GAACAACCTAAATGCTTAAATCTAATTTAGATGAGCATCAACCAAGCTAATACGAT 540  
 QY 541 AAAACGACTTTGATTAATGAACACCAAAATTTAGTTGAAGCATCAAAAGCTAATAAAC 600  
 DB 541 AAAACGACTTTGATTAATGAACACCAAAATTTAGTTGAAGCATCAAAAGCTAATAAAC 600  
 QY 601 ACTTGAACAACGCTTACTTAACCTTGAAGCTTGTATCAACGCTTATTAATGAAT 660  
 DB 601 ACTTGAACAACGCTTACTTAACCTTGAAGCTTGTATCAACGCTTATTAATGAAT 660  
 QY 661 CGCATATATTTAGTGAATCTATACATTAAGCTAGTATTTAATTAACCTAATAACCTAGAT 720  
 DB 661 CGCATATATTTAGTGAATCTATACATTAAGCTAGTATTTAATTAACCTAATAACCTAGAT 720  
 QY 721 CCAGTAAATGGGGGAAACGCTTTAGATTTCTAATGATTAATCAAGTTAATCGAATAT 780  
 DB 721 CCAGTAAATGGGGGAAATGCTTTAGATTTCTAATGATTAATCAAGTTAATCGAATAT 780  
 QY 781 AATATAGCTTATCAATTTAATGAACAAAGACTAATGCTGATGATTAACACCGCTCAT 840  
 DB 781 AATATAGCTTATCAATTTAATGAACAAAGACTAATGCTGATGATTAACACCGCTCAT 840  
 QY 841 TTTATTTAAAAAGTATCAAAATATATGAACAAAGTTTGTAGGAGCTTTTCAAAACGCT 900  
 DB 841 TTTATTTAAAAAGTATCAAAATATATGAACAAAGTTTGTAGGAGCTTTTCAAAACGCT 900  
 QY 901 AATGTTCAACCTTCAACCTCAAGTTTGTGCTTTTAAAGCTGATGATTAACACCGCTCAT 960  
 DB 901 AATGTTCAACCTTCAACCTCAAGTTTGTGCTTTTAAAGCTGATGATTAACACCGCTCAT 960  
 QY 961 TATTAATATGCAAGAGAGCGTTTGAATGATGAACTTCAAGTAGAATTC 1015  
 DB 961 TATTAATATGCAAGAGAGCGTTTGAATGATGAACTTCAAGTAGAATTC 1015  
 RESULT 9  
 565869 1015 bp DNA linear BCT 16-MAR-2000  
 565869  
 DEFINITION TM-1=29 kDa polypeptide [Mycoplasma gallisepticum, S6, ATCC 15302, Genomic, 1015 nt].  
 ACCESSION 565869  
 VERSION 565869.1 GI:425376  
 KEYWORDS  
 SOURCE Mycoplasma gallisepticum  
 ORGANISM Mycoplasma gallisepticum  
 Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 REFERENCE 1 (bases 1 to 1015)

AUTHORS Saito, S., Fujisawa, A., Ohkawa, S., Nishimura, N., Abe, T., Kodama, K., Kamogawa, K., Aoyama, S., Iritani, Y., and Hayashi, Y.  
 TITLE Cloning and DNA sequence of a 29 kilodalton polypeptide gene of Mycoplasma gallisepticum as a possible protective antigen  
 JOURNAL Vaccine 11 (10), 1061-1066 (1993)  
 MEDLINE 94025893  
 PUBMED 8212828  
 REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gi395 138145] from the original journal article. This sequence comes from Fig. 2.  
 FEATURES  
 SOURCE  
 1. 1015  
 /organism="Mycoplasma gallisepticum"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:2096"  
 202. .>1015  
 /gene="TM-1"  
 /note="29 kDa polypeptide; This sequence comes from Fig. 2; conceptual translation presented here differs from translation in publication"  
 /codon\_start=1  
 /transl\_table=4  
 /product="TM-1"  
 /protein\_id="AAB28343.2"  
 /db\_xref="GI:7249262"  
 /translation="MNRKRIILKTLISLGTSPISLIGSGMSITKDKANPNQOTOLQAARWELDLINAKRTLASLDYDAKEBASISSVSESELYNNNLNATLEOLKAKTNLESAINQANDKTFEDNEHPULVEAYALKTLEORATNLEGLASTAVNOIRNNLVLDYNNASLITTLTLDPLNGGMLDSNEITTVNINNTLSTIEQKNDALNSFTKVIONNESFVGFFTNANVOPSVSVASADVTVPNYKXARTVNWGDEPSSRI"  
 BASE COUNT 401 a 161 c 138 g 315 t  
 ORIGIN  
 Query Match 75.5%; Score 986.2; DB 1; Length 1015;  
 Best Local Similarity 98.2%; Pred. No. 5.1e-134;  
 Matches 997; Conservative 0; Mismatches 18; Indels 0; Gaps 0;  
 QY 1 AAAAATCATGAGTTGTTATCTGATATCTTTGCTTAAAAAACAACAAATCTTCTAACAA 60  
 DB 1 AAAAATCATGAGTTGTTATCTGATATCTTTGCTTAAAAAACAACAAATCTTCTAACAA 60  
 QY 61 AATCTTAATTAATTAAGCGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120  
 DB 61 AATCTTAATTAATTAAGCGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120  
 QY 121 AACCAAAATCTCTAGTATATAACCGTTATTTATTTATTTAGTCATCTTTAGAT 180  
 DB 121 AACCAAAATCTCTAGTATATAACCGTTATTTATTTATTTAGTCATCTTTAGAT 180  
 QY 181 ATTAATATATCTTATATTTCTTGAATAAGAAAGATCATCTTAAAGACTTTAGTTG 240  
 DB 181 ATTAATATATCTTATATTTCTTGAATAAGAAAGATCATCTTAAAGACTTTAGTTG 240  
 QY 241 TTAGGTACAAACCTCTTCTTGAATTTAGGATTTCTAGCTGATCTTATTAATAAAA 300  
 DB 241 TTAGGTACAAACCTCTTCTTGAATTTAGGATTTCTAGCTGATCTTATTAATAAAA 300  
 QY 301 GATGCAAAACCAAAATATGCGCAAAACCAATTTAGAGAGCGGATGAGATTAACAGAT 360  
 DB 301 GATGCAAAACCAAAATATGCGCAAAACCAATTTAGAGAGCGGATGAGATTAACAGAT 360  
 QY 361 CTATATCATGCTTAAAGCGATGACATTAGCTTACATCAAGACTATGCCAAGTTGAAGCT 420  
 DB 361 CTATATCATGCTTAAAGCGATGACATTAGCTTACATCAAGACTATGCCAAGTTGAAGCT 420  
 QY 421 AGTTATCATGCTTATAGTGAAGCTGAAACAGTTAACATTAACCTTAAATGCAACATTA 480  
 DB 421 AGTTATCATGCTTATAGTGAAGCTGAAACAGTTAACATTAACCTTAAATGCAACATTA 480  
 QY 481 GAACAACCTAAATGCTTAAATCTAATTTAGATGAGCATCAACCAAGCTAATACGAT 540

LOCUS	AX665187	1082 bp	DNA	linear	PAT 26-MAR-2003
DEFINITION	Sequence 24 from Patent Epi275716.				
ACCESSION	AX665187				
VERSION	AX665187.1	GI:29290312			
KEYWORDS					
SOURCE	Mycoplasma gallisepticum				
ORGANISM	Mycoplasma gallisepticum				
REFERENCE	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.				
AUTHORS	1 Okuda, T., Satou, S., Dorsey, K.M. and Tsuzaki, Y.				
TITLE	Modified dna molecule, recombinant containing the same thing, and				
JOURNAL	Patent: EP 1275716-A 24 15-JAN-2003;				
FEATURES	Zeon Corporation (JP)				
source	Location/Qualifiers				
	1..1082				
	/organism="Mycoplasma gallisepticum"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:2096"				
	/note="Modified TTM-1 portion (downstream of Bgl1) of				
	pNZ40K-S"				
BASE COUNT	403 a	184 c	183 g	312 t	
ORIGIN					

[illegible]

QY	475	ACATTGGAACACTAAAAAGGCTAAACCTAATTATGATACGACCTCAACCAAGCTAAT	534
Db	121	ACATTGACACACTAAAAATGCTAAACCTAATTGTAATCAGCCATCAACCAAGCTAAT	180
QY	535	ACGGATAAAAAGACTTTGGTAAATGAAACCCAAATTAGTGAAGCTATCAAAAGCACTA	594
Db	181	ACGGATAAAAAGACTTTGGTAAATGAAACCCAAATTAGTGAAGCTATCAAAAGCACTA	240
QY	595	AAAAACCTTTAAGAACACGCTGCTACTAACCTGAAGGTTTGCATCAACTGCTTAAT	654
Db	241	AAAAACCTTTAAGAACACGCTGCTACTAACCTGAAGGTTTGCATCAACTGCTTAAT	300
QY	655	CAAAATGCAATATATTAGTGAATCTATACAAATAAAGCTAGAGTTTAATTAACATAAACA	714
Db	301	CAAAATGCAATATATTAGTGAATCTATACAAATAAAGCTAGAGTTTAATTAACATAAACA	360
QY	715	CTAGATCCACTAAATGGGGGAACGCTTTTGATTTCTATAGAGTTACTAAGATTAATCGG	774
Db	361	CTAGATCCACTAAATGGGGGAACGCTTTTGATTTCTATAGAGTTACTAAGATTAATCGG	420
QY	775	AATTTAAATAAAGCTATCAACTATTAATGAACAAAGAAGCTAATGCTGATGATTAAT	834
Db	421	AATATCCAGAAATAGTTATCAACTATTAATGAACAAAGAAGCTAATGCTGATGATTAAT	480
QY	835	AATAGTTTATTAATAAAAAAGTATCAAAATATGAAACAAGTTTGTAGGAGCTTTTACA	894
Db	481	AATAGTTTATTAATAAAAAAGTATCAAAATATGAAACAAGTTTGTAGGAGCTTTTACA	540
QY	895	AACCTTAATGTTCAACTCTCAAACTAAGTTTTGCTTTTATGCTGATGATTAACACC	954
Db	541	AACCTTAATGTTCAACTCTCAAGCAATAGTTTGTCTTTATGCTGATGATTAACACC	600
QY	955	GTCATTAATAATATGACAGAAAGACCGTTTGGATGATGATGAACTTCAGTAGAATT	1014
Db	601	GTCATTAATAATATGACAGAAAGACCGTTTGGATGATGATGAACTTCAGTAGAATT	660
QY	1015	CTTGAAACACGAATAGTATCACAGATGTTCTTGATTTATTAATTTAGCTGGAAACAAC	1074
Db	661	CTTGAAACACGAATAGTATCACAGATGTTCTTGATTTATTAATTTAGCTGGAAACAAC	720
QY	1075	ACGAAGTACCAATTAGTTTATAGAACATATGATGCTCACTCAACTGCTTATTAATTTCCCT	1134
Db	721	ACGAAGTACCAATTAGTTTATAGAACATATGATGCTCACTCAACTGCTTATTAATTTCCCT	780
QY	1135	TATAAGTTGTAAAGACAGCTGATGCTAATAACGTTGGATTACAATACAAAATTAATAT	1194
Db	781	TATAAGTTGTAAAGACAGCTGATGCTAATAACGTTGGATTACAATACAAAATTAATAT	840
QY	1195	GGAAATGTTCAACAAGTTGAATTTGCCCTTCAACTAGTGAATAATAATACACAGCTAAT	1254
Db	841	GGAAATGTTCAACAAGTTGAATTTGCCCTTCAACTAGTGAATAATAATACACAGCTAAT	900
QY	1255	CCAACTCCAGAGGTTGATGAGATTAATAATTGCTAAATCGTTTTATCAGGTT	1306
Db	901	CCAACTCCAGAGGTTGATGAGATTAATAATTGCTAAATCGTTTTATCAGGTT	952

LOCUS	853 bp	DNA	linear	PAT 29-SEP-1997
RESULT 11				
E02348	E02348			
DEFINITION	DNA sequence coding for TMG-1.			
ACCESSION	E02348			
VERSION	E02348.1	GI:2170583		
KEYWORDS	JP 1990111795-A/7.			
SOURCE	Mycoplasma gallisepticum			
ORGANISM	Mycoplasma gallisepticum			
REFERENCE	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.			
AUTHORS	1 (bases 1 to 853)			
	Kodama, K., Saito, S., Yanagida, N., Kamogawa, K., Iritani, K. and			
	Aoyama, S.			
TITLE	FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE			
JOURNAL	THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME			
	Patent: JP 1990111795-A 7 24-APR-1990.			

COMMENT  
NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
OS Mycoplasma gallisepticum  
PN JP 1990111795-A/7  
PD 24-APR-1990  
PF 02-JUN-1989 JP 1989140283  
PR 02-JUN-1988 JP 88P 136343  
PI KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,  
PI IRIYANI KOICHI, AOYAMA SHIGEMI  
PC C07K13/00, C07K7/06, C07K15/04, C12N1/21, C12N15/31//A61K37/02, PC  
C12P21/02, PC  
G01N3/569, (C12N1/21, C12R1:19), (C12N15/31, C12R1:35), (C12P21/02, PC  
C12R1:19),  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: strain=56;  
FH key Location/Qualifiers  
FT 5'UTR 1..39  
FT CDS 40..825  
FT FT /product='MG-1',  
FT 826..853.  
FEATURES  
source  
1..853  
/organism="Mycoplasma gallisepticum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2096"  
BASE COUNT 329 a 138 c 128 g 258 t  
ORIGIN  
Query Match 63.1%; Score 824.2; DB 6; Length 853;  
Best Local Similarity 97.9%; Pred. No. 1.8e-110; Indels 0; Gaps 0;  
Matches 835; Conservative 0; Mismatches 18;  
QY 163 TTAGTCATCTTTTAAAGATATAATATCTTAATATCTATGATAGAAAAAGATATC 222  
DB 1 TTAGTCATCTTTTAAAGATATAATATCTTAATATCTATGATAGAAAAAGATATC 60  
QY 223 TTTAAGATCTTTAGTTGTTAGTACCAATCCCTTTCTTAGCATTTGGATTTCTAGCTGT 282  
DB 61 TTTAAGATCTTTAGTTGTTAGTACCAATCCCTTTCTTAGCATTTGGATTTCTAGCTGT 120  
QY 283 ATGTCTATTACTTAAAGATGCAAAACCAATATAGGCCAAACCAATTTGAAGCAGC 342  
DB 121 ATGTCTATTACTTAAAGATGCAAAACCAATATAGGCCAAACCAATTTGAAGCAGC 180  
QY 343 CGAATGAGTTAAACAGATCTATCAATGCTAAAGCATGATGATCTTCACTACAGAC 402  
DB 181 CGAATGAGTTAAACAGATCTATCAATGCTAAAGCATGATGATCTTCACTACAGAC 240  
QY 403 TATGCCAAGATTGAAGCTTATCATCTGCTTATAGTGAAGTGAACAGTTAACAT 462  
DB 241 TATGCCAAGATTGAAGCTTATCATCTGCTTATAGTGAAGTGAACAGTTAACAT 300  
QY 463 AACCTTATGCAACCTTAGAACAACTAAATGCTAAATGCTAAATTTGATCGCCATC 522  
DB 301 AACCTTATGCAACCTTAGAACAACTAAATGCTAAATGCTAAATTTGATCGCCATC 360  
QY 523 AACCAAGCTAATACGATTAACGATTTTGTATATGAACCCCAATTTAGTTGAACA 582  
DB 361 AACCAAGCTAATACGATTAACGATTTTGTATATGAACCCCAATTTAGTTGAACA 420  
QY 583 TACCAAGCATTAACCACTTTAGAACAGTGTCTACTAACCTTTGAAGTTTGTATCA 642  
DB 421 TACCAAGCATTAACCACTTTAGAACAGTGTCTACTAACCTTTGAAGTTTGTATCA 480  
QY 643 ACTGCTTATTAATCAATTCGCAATATTTAGTGATCTATCAATTAAGAGTATAGTTA 702  
DB 481 ACTGCTTATTAATCAATTCGCAATATTTAGTGATCTATCAATTAAGAGTATAGTTA 540  
QY 703 ATTAATAAAGCATAGATCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACT 762

DB 541 ATTAATAAAGCATAGATCACTAAATGGGGGAACGCTTTTAGATTCTAATGAGATTACT 600  
QY 763 ACAGTTAATGGAATATTAATTAATACGTTATCACTTTAATGAACAAAGACTTAAGCT 822  
DB 601 ACAGTTAATGGAATATTAATTAATACGTTATCACTTTAATGAACAAAGACTTAAGCT 660  
QY 823 GATCATTAATCTAATTAATTTTATTAATAAAGTGAATTCAAATTAATGAACAAAGTTTGT 882  
DB 661 GATCATTAATCTAATTAATTTTATTAATAAAGTGAATTCAAATTAATGAACAAAGTTTGT 720  
QY 883 GGGACTTTTACAAACGCTAATGTTCAACCTTCAACTAGTCTTTTGTCTTTAGTCT 942  
DB 721 GGGACTTTTACAAACGCTAATGTTCAACCTTCAACTAGTCTTTTGTCTTTAGTCT 780  
QY 943 GATCTAACACCCGCTCAATTTAATTAATGCAAGAGACCCGTTGGATGCTGAAGAACT 1002  
DB 781 GATCTAACACCCGCTCAATTTAATTAATGCAAGAGACCCGTTGGATGCTGAAGAACT 840  
QY 1003 TCAAGTAGAATTC 1015  
DB 841 TCAAGTAGAATTC 853  
RESULT 12  
E02342 708 bp DNA linear PAT 29-SEP-1997  
LOCUS  
DEFINITION DNA sequence coding for MG-1.  
ACCESSION E02342  
VERSION E02342.1 GI:2170577  
KEYWORDS JP 1990111795-A/1.  
SOURCE Mycoplasma gallisepticum  
ORGANISM Mycoplasma gallisepticum  
REFERENCE 1 (bases 1 to 708)  
AUTHORS Kodama,K., Saito,S., Yanagida,N., Kamogawa,K., Iriyama,K. and Aoyama,S.  
TITLE FOWL MYCOPLASMA ANTIGEN, RECOMBINANT VECTOR CONTAINING GENE THEREOF, DIAGNOSTICUM AND VACCINE USING THE SAME  
JOURNAL Patent: JP 1990111795-A 1 24-APR-1990;  
COMMENT NIPPON ZEON CO LTD, SHIONOGI & CO LTD  
OS Mycoplasma gallisepticum  
PN JP 1990111795-A/1  
PD 24-APR-1990  
PF 02-JUN-1989 JP 1989140283  
PR 02-JUN-1988 JP 88P 136343  
PI KODAMA KAZUMI, SAITO SHUJI, YANAGIDA NOBORU, KAMOGAWA KOICHI,  
PI IRIYANI KOICHI, AOYAMA SHIGEMI  
PC C07K13/00, C07K7/06, C07K15/04, C12N1/21, C12N15/31//A61K37/02, PC  
C12P21/02, PC  
G01N3/569, (C12N1/21, C12R1:19), (C12N15/31, C12R1:35), (C12P21/02, PC  
C12R1:19),  
CC strandedness: Double;  
CC topology: Linear;  
CC \*source: strain=56;  
FH key Location/Qualifiers  
FT 5'UTR 1..708  
FT CDS 1..708  
FT FT /product='MG-1',  
FT 709..762.  
FEATURES  
source  
1..708  
/organism="Mycoplasma gallisepticum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:2096"  
BASE COUNT 278 a 119 c 107 g 204 t  
ORIGIN  
Query Match 52.1%; Score 679.8; DB 6; Length 708;  
Best Local Similarity 97.6%; Pred. No. 1.7e-89;  
Matches 690; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 280 TGTATGCTATTACTAATAAAGATGCAAAACCAATTAATGCGCAACCAATTAAGAAACA 339







Db 258967 TTTCAAACTGAAACCTGAAAAAACAATGCTAATACCTTACGTAACAAGTTTGTAAAG 259026  
 Qy 851 AAGTATTCMAAATATGAAACAAAGTTTGTGAGACTTTTACAAACCTTAATGTTCAAC 910  
 Db 259027 AGTTTGTAGTAAATAATTAATACAGAAATGATAC--AACTAATATGAGAAACAC 259083  
 Qy 911 CTTCAACTACAGTTTGTGCTTTAGTGCATGATGAACCCCGTCAATTAT----- 963  
 Db 259084 CTGGAATATACAGCTTTGTGTGTTAGTGTAAATGCAACAATAATATGAGATTCCTA 259143  
 Qy 964 -----AATATGCAAGAAAGACCGTTGGAATGATGAACCTTCAAGTAATTTCTTG 1018  
 Db 259144 ACTGGAATTTGCAACAAAGAAAGTTGAACATGATGAACGTTGAACATGCTAATTT 259203  
 Qy 1019 CAACACGAATAGTA-----TCAGATGTTTCTTGATTAATAGTTAGCTG 1066  
 Db 259204 CTAGCACTTCTGATTAATTTCTTCTACTTAACAGAAATCATGAATTTATAGTTATCAG 259263  
 Qy 1067 GAACAACAGCAAGTCAATTAATTAAGTAAAGCACTATGTCATCAACTGCTGTTATTTAT 1126  
 Db 259264 GAGCAGCTCTAATATACAGTTTAACTTAATTAATTAAGCTGCTCACTGTTATTTAT 259323  
 Qy 1127 ATTTCCCTTAATAAGTTGTTAAAGCAGCTGATGCTAATAAGCTGATTAACAAT 1186  
 Db 259324 ATTTCCCTTAATAAGTTGTTAAACAGAGTGATGAATAATGATGCTTACATATACCT 259383  
 Qy 1187 TAAATATGAATGTTTCAACAGTGAAGTTTGCCACTTCAACTGATGAATAATATCTA 1246  
 Db 259384 TAAATGATGATGACAGAAAGTTAAGTTTCTCAACTGTAATAACTAGTGTAGTG 259443  
 Qy 1247 CAGCTAATCCACTCCAGCAGTTGATGAGTTAAAGTCT 1287  
 Db 259444 CAAATCTTCTGCTGATGAATGAATATCAACAGAAAGTCT 259484

RESULT 14  
 AF210770 4568 bp DNA linear BCT 02-FEB-2000  
 LOCUS Mycoplasma gallisepticum pmga-3-like protein 9.1 gene, partial cds;  
 DEFINITION pmga-like protein 9.2 gene, complete cds; and pmga-like protein 9.3  
 gene, partial cds.  
 ACCESSION AF210770.1 GI:6851355  
 VERSION AF210770  
 KEYWORDS Mycoplasma gallisepticum  
 SOURCE Mycoplasma gallisepticum  
 ORGANISM Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 REFERENCE 1 (bases 1 to 4568)  
 AUTHORS Phair,G.T., Branton,S.L., Hanson,L.A., Minton,F.C., Lotz,B.D.,  
 May,J.D. and Hughtlett,M.B.  
 TITLE A novel pmga-like gene from the F-strain (vaccine strain) of  
 Mycoplasma gallisepticum  
 JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 4568)  
 AUTHORS Phair,G.T., Branton,S.L., Hanson,L.A., Minton,F.C., Lotz,B.D.,  
 May,J.D. and Hughtlett,M.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-DEC-1999) College of Veterinary Medicine, Mississippi  
 State University, Box 9825, MS 39762, USA  
 FEATURES  
 SOURCE  
 1..4568  
 /organism="Mycoplasma gallisepticum"  
 /mol\_type="genomic DNA"  
 /strain="F"  
 /db\_xref="taxon:2096"  
 <1..637  
 /note="adhesin"  
 /codon\_start=2  
 /transl\_table=4  
 /product="pmga-like protein 9.1"  
 /protein\_id="AAF29523.1"  
 /db\_xref="GI:6851356"

repeat\_region  
 CDS  
 /translation="VYNDIFSNILDKODEPNSVTYNLKGYSIADHSYFYOFSSN  
 RNESNPYLVGFIGGHRNRNINSNTNNVNASPVQTSNRLLIIVNPKQCY  
 IKGSLYISNNRNLKPTTATANNSTLFTVKGANNSTGTGRTANNNDIENS  
 GSSSSG  
 OPENESTKILINKLNGKLVITSVMDNKNPAGYIGNLFTLLMSPTEMAKK"  
 787..831  
 /note="microsatellite"  
 /rpt\_type=tandem  
 /rpt\_unit=gaa  
 /rpt\_count=2982  
 /note="adhesin"  
 /codon\_start=1  
 /transl\_table=4  
 /product="pmga-like protein 9.2"  
 /protein\_id="AAF29524.1"  
 /db\_xref="GI:6851357"  
 /translation="MKRKNIILKRVSLIGISFVMLAAASCTATPTTPEPKPDPMP  
 NPSGMDNGDTPNGDGMENSAQOLAARKEBLSLILATOSNNLAKVADYNINOTL  
 TAAVTTASTSDNTSVLEQVKSATSTLAAIDTAASKTSFDEKPELITKAYALKE  
 TLAKNEEVLISGLTDSNFATIKMLTALYOSGDFVATLDPVGNAPQIADITKADK  
 IADAVSKLETWKTNNATLATSFEKVELYKXKLGIDITNNRQPGNYSFVGSVANTN  
 MEIIPNNPQARVMTSDNGRTSLISTSDNSSTLFEVMIYSLGAGTSLFENY  
 GSESTGLYFFPKLVKSGDENVALQYTLASGAEVNFAPYKTSVSAUSSGDSNOT  
 ESAEETMPVSDINPAVTSIDINIAKLTLSNLFKSGNTIEFSVPEPSKVAPMIGM  
 YLTSNLANEAKVNSIFGVNDNSEASTVYVDILKYSLATWSTVTRFANUTISM  
 PENATYLVGFIGQLARTVGSIPNNRNPINNNENRFTLYVNAPKAGDIHISGSY  
 LTRNTGLKLTVDITDKNNSITITSGKNNTLGHFPTSKNNSNGNDGSVENNKA  
 SILKRGKNTVIAGTQDGKNAFYIGNLFTLINSSTNASQSSST"  
 3147..3182  
 /note="microsatellite"  
 /rpt\_type=tandem  
 /rpt\_unit=gaa  
 /rpt\_count=2982  
 /note="adhesin"  
 /codon\_start=1  
 /transl\_table=4  
 /product="pmga-like protein 9.3"  
 /protein\_id="AAF29525.1"  
 /db\_xref="GI:6851358"  
 /translation="MKRKNIILKRVSLIGISFVMLAAASCTAATPTPEPKPDPMP  
 KPDMPNPSGMDNGDTPNGDGMENSAQOLAARKEBLSLILATOSNNLAKVADYNINOTL  
 AKIONTLTAATTAATGRTSSATLEKVNATSLTQATINTANSNKKPODSNLM  
 SYKMLMTLAKETITVTLDPRKSLIDIDINGSCGEISLVDHDDPAGYVAAAT  
 ITBEITKIEVISEKTIQDQKNNADQDNQSPFLDKTLENVEDAKMOPGANYSV  
 GYSVDVGTSGQETIIPNNPQARVMTSDNGRTSLISTSDNSSTLFEVMIYSLGAGTSLFENY  
 SLATGAKYTLFTFYGSPGWLFFPKLVKANDVGLQYKANSNETLPIIFEGTT  
 TNGPAATVENINNAKRVLTGLAFKNT"

BASE COUNT 1726 a 706 c 760 g 1376 t  
 ORIGIN  
 Query Match 19.1%; Score 249; DB 1; Length 4568;  
 Best Local Similarity 55.2%; Pred. No. 3.3e-27;  
 Matches 586; Conservative 0; Mismatches 445; Indels 30; Gaps 4;

Qy 251 CATCTTTCTTACGATTGGAGTTTCTAGCTGATGCTAATTAATAAATGCAACC 310  
 Db 1097 CAACCCCTCTAGTGTATATGAATGGCGAGATCTAATCCAGAAATGACGAGAA 1156  
 Qy 311 CAATATATGCCCAACCAATTAGAGCGCGAATGAGATTAAACAGATCAATG 370  
 Db 1157 TGAAGATTTCTGCTCAACATTAAGACGCTGTAATAAAGATTAAAGCATTTATAGCA 1216  
 Qy 371 CTAAGCGATGACATTAGCTTCACTAACAAGCTATGCCAAGTTGAAGCTAGTTATCAT 430  
 Db 1217 CACAAAGCAGTAACCTTGCAAAATATGCAATTAACAAATTTGCAATCTTAACAG 1276  
 Qy 431 CTGCTATATGTAAGCTGAAACAGTTAAACAATCTTAATAGCAACATTAGAACACTAA 490  
 Db 1277 CTGCATACCAACTGTATTAATCTACATGATTAACAAGCTGCTTAAGAGCAAGTTA 1336  
 Qy 491 AATGCTTAAACCTAATTTGAATGACGCATCAACCAAGCTAATACGATTAACAGACT 550  
 Db 1337 AAGGCTACATCACTTACAGAGGATATGATGATGCTGCAAGTTCAAAACTAGTT 1396



```

Oy 551 TTGATATGACACCCAAATTAGTTGAAGCATACAAAGCACTAAAAACACTTAGAAC 610
Db 1397 TTGATGAAAAAATCTGATATTAATCAAGCATATTATGCTTTAAAGAAACATTAAAAA 1456
Oy 611 AACGCGTACTACCTTGAAGTTTGTGATCACTGCTTATATCAAAATTCGCATATT 670
Db 1457 ATGAAGAAACAGTTTAAAGTGTTAAGTCTGATTTCAATTTTGCACAAATCAAACTAAC 1516
Oy 671 TAGTGATCTATACAAATTAAGCTAGTACTAGTTTAAATTAATAAAACATAGTACCAATAG 730
Db 1517 TTACGCTTTATATATCAATCTGTTAAAGATTTTGTTTAAAGCAACATTAGTCTGTG---TA 1573
Oy 731 GGGGAGCCCTTTAGATTCTTAATGAGATTACTACAGTAATCGAATAATTAATAATACGT 790
Db 1574 GTGGAAGAGCTCCCAATGCTGATATTACAAAGCAGATTAAGATATTGCTGATGCTG 1633
Oy 791 TATCAACTATTTATGACAAAGAAAGTAAATGCTGATGATATCTATATGTTTATATAAA 850
Db 1634 TTTCAAACTTGAAACTGAAAAACAAATGCTTAATCTTACCTCAAGTTTGTGTTAAG 1693
Oy 851 AAGTATTCAAAATTAATGACAAAGTTTGTAGGAGCTTTTCAAAAGCTAATGTTCAAC 910
Db 1694 AGGTTTATGTTAAATTAATTAACAGGAATTGATAC---AACTAATTAATCGAGAACAC 1750
Oy 911 CTTCAACTACAGTTTGTGCTTTTGTAGTCTGATGATACACCCGCTCAATTAT----- 963
Db 1751 CTGGAATTTACAGCTTTGTTGTTAGTGTAAATGCAACAAATTAATATGATGATTCCTA 1810
Oy 964 -----AATATGCAAGAAAGACCGTTTGAATGATGATGAACCTCAAGTAAATTCCTG 1018
Db 1811 ACTGGAATTTGCGACAAAGAAAGTTTGAACCTAGATGAACCGTAACTAGTCTAATTT 1870
Oy 1019 CAACACAGCAATAGTA-----TCACAGATGTTCTTGAATTTATAGTTTAGCTG 1066
Db 1871 CTAGACATCTTGATTAATCTTCTACCTTAACAGAGTATCATGATTTATAGTTTATCAG 1930
Oy 1067 GAACAAACAGAAAGTACCAATTTAGTTTACCACTATAGTGCATCAACCTGTTATTTAT 1126
Db 1931 GAGCAGGATACCTAAATTAACAGCTTAATTAATTAACGCTCTCACTGTTATTTAT 1990
Oy 1127 ATTTCCCTTATTAAGTTGTTAAAGCACTGATGCTAATTAACCTGATTAACATCAAT 1186
Db 1991 ATTTCCCATATTAAGTTAGTTAAAGAAAGGATGAAGAAATTAATCTTACCTTACATTA 2050
Oy 1187 TAAATTAATGGAATGTTCAACAAGTTGATTTGCCACTTCAACTGATGCAATATATACGA 1246
Db 2051 TAAATTAATGTTAGTGAACAGAGTTGATGATTAAGTTGCT 1287
Oy 1247 CAGCTAATCCAACTCCAGCAGTTGATGATTAAGTTGCT 1287
Db 2111 CAGATTCTTCTGCTGATTTCAATTAATTAATCAGACAGAAAGTCT 2151

```

```

RESULT 15
LOCUS AE016969/c 301903 bp DNA linear BCT 09-JUN-2003
DEFINITION Mycoplasma gallisepticum strain R section 3 of 4 of the complete
genome.
ACCESSION AE016969 AE015450
VERSION AE016969.1 GI:31541483
KEYWORDS
SOURCE Mycoplasma gallisepticum R
ORGANISM Mycoplasma gallisepticum R
Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
REFERENCE
AUTHORS Gearty,S.J., Papadatos,L., Kutish,G., Gorton,T.S., Mahairas,G.,
Swartzell,S., Madan,A., Nguyen,D.K., Markham,P., Browning,G.,
Kamal,M. and Liao,X.
TITLE The complete genome sequence of the avian pathogen Mycoplasma
gallisepticum strain R
JOURNAL Microbiology (2003) In press
REFERENCE 2 (bases 1 to 301903)
AUTHORS Gearty,S.J., Papadatos,L., Kutish,G., Mahairas,G., Swartzell,S.,

```

## TITLE

JOURNML

Madan,A., Nguyen,D.K., Gorton,T.S., Markham,P., Browning,G.,  
Mustafa,K. and Liao,X.  
Submitted (17-OCT-2002) Department of Pathobiology and Veterinary  
Sciences, and Center of Excellence for Vaccine Research, The  
University of Connecticut, 61 North Eagleville Road U-89, Storrs, CT  
06269-3089, USA

## FEATURES

Location/Qualifiers

1..301903

/organism="Mycoplasma gallisepticum R"

/mol\_type="genomic DNA"

/strain="R"

/db\_xref="taxon:233150"

/gene

complement(17..754)

/CDS

/locus\_tag="MGA\_0090"

complement(177..754)

/locus\_tag="MGA\_0090"

/note="MGA\_0090; [R] COG1881 Phospholipid-binding protein; MGR\_456"

/codon\_start=1

/transl\_table=4

/product="predicted phospholipid-binding protein"

/protein\_id="AAP56784.1"

/db\_xref="GI:31541484"

/translation="MRIMKXYSIGNSKIMSRYVADGYLALHGGHKNENQTPPSF

SLDSWEKIPNAKSYAVIVEDPAAVIGSFIMIVIKTHNIAENOSYLPHNKM

ESKOSAGENDLMQGHNSVVKTLVANNKTNKDLGILPEGFTSLSENSIMYGCYP

PNDHTYTVVVYGLDVDASELTYHKHRTQKYLHNKRYVYGDPMQALHKHVVGHHTLH

FKRKVG"

complement(767..1396)

/CDS

/locus\_tag="MGA\_0091"

complement(767..1396)

/locus\_tag="MGA\_0091"

/note="MGA\_0091; [R] COG1881 Phospholipid-binding protein; MGR\_457"

/codon\_start=1

/transl\_table=4

/product="predicted phospholipid-binding protein"

/protein\_id="AAP56785.1"

/db\_xref="GI:31541485"

/translation="MKRTISYIFNLATYFNDYILMTIMDMKKHAIKSNCFETGDS

VYLSKSTVKNLNKNNYESLDLTWDKIEGAKSYAVVAVYASRVIGSFIMVAVNKS

NKLAVYANINNKSIYVQGINSTAQCTSSKGVILIECVGAFKNTAEASDYLPLPPD

DTHLYTLRVYGLDVHDVLDLPKGFSLTDLNQKTIIRHTVGEHLHFMYKEN"

complement(1437..2306)

/gene="ara1"

/CDS

/locus\_tag="MGA\_0093"

complement(1437..2306)

/gene="ara1"

/locus\_tag="MGA\_0093"

/note="MGA\_0093; ARA1 [R] COG0656 Aldo/keto reductases

related to dikeetoglucuronate reductase; MGR\_458"

/codon\_start=1

/transl\_table=4

/product="ARA1"

/protein\_id="AAP56786.1"

/db\_xref="GI:31541486"

/translation="MYLTTRKIKMKYKYLANNYKMPISGRGTYKLEDKQTVAVKT

ALBVGYRHLDCAEITENQSVSGIKLSGVNRKIFITSKIANDKGESEKAPFKI

LKLDLDLEVLDDLHPIDFGKFKDQWQVNAQTWAKMAEFLERIKATIGSNFLVH

IEALKSKAKILPMVNOLEFHFGYLOPEIVBVCQKNINIVQAMSPFSOSKYLENTLVO

LAKKYETVAQILINWIIIOKDIYVLPKSKTPERIRSNFVDFNFGINQADLKVIDLND

CGKGLHPDQVDF"

2541..2930

/locus\_tag="MGA\_0094"

/CDS

/locus\_tag="MGA\_0094"

/note="MGA\_0094; MGR\_459"

/codon\_start=1

/transl\_table=4

/product="conserved hypothetical"

/protein\_id="AAP56787.1"

/db\_xref="GI:31541487"

**gene**

/translation="MINLLIPFWFNKSKOLLISKPRITYINLKSNNISKLVRVDLSV  
ILIOLEYMMATFRKKRRLIKKDIDLIDSEWNMDLKVIISFKSLHRWDRDEIDYFTSR  
IINQNKNIPLFINRRRPDOLLRARCK"  
complement(2997..3440)  
/locus\_tag="MGA\_0096"  
/locus\_tag="MGA\_0096"  
/note="MGA\_0096; MGR\_460"  
/codon\_start=1  
/transl\_table=4  
/product="unique hypothetical"  
/protein\_id="AAP56788.1"  
/db\_xref="GI:31541488"  
/translation="MKKKRKVASITKMNLKAELKTMTKEELKERLNVSDDVSQSLLL  
ALGEPKIAYNDLIANTYTKLVGESEOILDIMNLANSPFGIEKEIKOEDIVNDPPASDVVR  
KMAEMEKLOSNYGDAIPFHDHYGAESDEDKEDIAPKDKKNLS"

**CDS**

complement(3397..3609)  
/locus\_tag="MGA\_0097"  
complement(3397..  
/locus\_tag="MGA\_0097")  
/locus\_tag="MGA\_0097"  
/note="MGA\_0097; MGR\_461"  
/codon\_start=1  
/transl\_table=4  
/product="unique hypothetical"  
/protein\_id="AAP56789.1"  
/db\_xref="GI:31541489"  
/translation="MSSPFLGVMIIGFFIHIERNNNSODLAYVTNLPIKKNYFPFIKITK  
VOIKOLIAKALKIKEISSINDOVK"

**gene**

complement(3483..3872)  
/locus\_tag="MGA\_0099"  
complement(3483..  
/locus\_tag="MGA\_0099")  
/locus\_tag="MGA\_0099"  
/note="MGA\_0099 unique hypothetical; MGR\_462"  
/codon\_start=1  
/transl\_table=4  
/product="nlike hyphtetical"  
/protein\_id="AAP56790.1"  
/db\_xref="GI:31541490"  
/translation="MITRMRSNMNPSSOIIDEINNISKVKOPMLIIVGFIGIIVAEG  
IGVISLPDSDNNIRSLKTNMDRTRENSISINSFVVADRYYHFMALMWLGFSY  
TSKMEILLKWLIKDIYSWKNIIFPK"

**CDS**

complement(3830..5323)  
/gene="pepc"  
/locus\_tag="MGA\_0100"  
complement(3830..  
/gene="pepc")  
/locus\_tag="MGA\_0100"  
/note="MGA\_0100; PepC [E] COG3579 Aminopeptidase C;  
MGR\_463"  
/codon\_start=1  
/transl\_table=4  
/product="PepC"  
/protein\_id="AAP56791.1"  
/db\_xref="GI:31541491"  
/translation="WMARKLELDLSKOLANNLSRSKENKOLTNTLTWTGWXSVSNGYN  
FSKNSISEFDLASOPTINOFOGTGCWI FAGINILRYHLAKELINIDLESOSYLAWPW  
DKFEANVALETVELTRDXKWDRIETFLKGNADVADGGOWTMLNLKKGLGSXSTM  
PTPASASNOCUNLYNILINKNOAASKIIDDNOKDPVALINKIKTALADEVFLSIISYG  
BLIPTPEFSYTVCIAKSDASHSESHPNYLPKKVVFENTPTLTPQQKFPAKVVEQGSGV  
SVYNVPANKPNQTFVSFYLNNVIESDELIAHTVNEQLFSYLTIKOLAVGPWFEC  
EVKNYVLINAKTFPDOSDFDYOTLFENIDSLSKNDOLDYWLSINHAMLITGDVLDLE  
KRNQIDOENAOINKANKTVKOAYOVEAHNDOLTIHKWKIENSWCODIGHAGYFVIDSOS  
YKRKTUYQAIVNROLFNOTLEVQLKTKERDKRVRLLEPMPIDTLAK"

**gene**

complement(5320..6972)  
/gene="pepc"  
/locus\_tag="MGA\_0101"  
complement(5320..  
/gene="pepc")  
/locus\_tag="MGA\_0101"  
/note="MGA\_0101; PepC [E] COG3579 Aminopeptidase C;  
MGR\_464"  
/codon\_start=1  
/transl\_table=4

[illegible]

```

Qy 803 ATGAACAAAGACTAATGCTGATGCTATTAATAGTTTATTAATAAAGTATTCAA 862
Db 220060 AACATGAAAAATATGCAAAACATGATATCTGATAGTTTTTAAGAAACCTTAGATA 220001
Qy 863 ATATGAACAAAGTTTGTAGGACTTTTACAAAGCTAATGTTCAACCTTCAACTACA 922
Db 220000 AAGAACAAATTAACAAGTAATGGGAATTACA-----CAACCTGCAAAATTACA 219953
Qy 923 GTTTGTGCTTTTGTAGTCTGATGTAACACCCGTCATTTATTAATATGCAAGAAAGACCG 982
Db 219952 GTTTGTGCAATATATCAAGTATTAATCTTACTTATTAATTTTGTCTAAGAGATAG 219893
Qy 983 TTTGGAATGATGATGAACCTTCAGTGAATTTCTTG-----AAACAGATAGTA 1033
Db 219892 TTTGAAACCAACAAGAGTGGTCAAGTACCTATGTCATTAAGAAATCAAGGTGATT 219833
Qy 1034 TCACAGATGTTTCTTGATTTATATGTTAGTGGAAACAAACAGAGTACCAATTTAGTT 1093
Db 219832 TAACGAGATATCATGAATTTATAGTTAGCTGGAAATGAACTAATATAGTTTACTT 219773
Qy 1094 TTAGCACTATGATGTCATCACTGCTTATTTATATTTTCCCTTATATAGTTGTTAAGCAG 1153
Db 219772 TTGCTAATTAATGCAACACTACAGATACTTATATTTCCATACAAAGTCAAGTTAATCA 219713
Qy 1154 CTGATGCTAATTAACGTTGATTTACATAACAATTAAATGAATGTAATGTTCAACAAGTTG 1213
Db 219712 GCG-----ATAATGTTGATTAACAATTAAGTTAATGATGCTAGTCCGGTATCTATCG 219659
Qy 1214 AGTTGGCACTTCAACTA 1231
Db 219658 ATTTAATTACTCAGCAA 219641

```

Search completed: August 25, 2003, 04:51:42  
 Job time : 5038 secs

THIS PAGE BLANK (USPTO)

293 YARRTWNGDEPSRILANTNSITDVSWISLA-----GTNTKYQFSFNVGP 340  
Db 724 KAQDALKAOVTSKORVANVTSIOQTANVELNTANGOLGIDDENATKOTOKTADAEOK 783  
Qy 341 STGYLPPFKLVKADA-----NNVGLQVYKLN-----NGNVQVEFATSTS 381  
Db 784 KTAIV-----DQAVAAKAILNKQTSNSDKAAVDRALQOVSTKHALNGDALAEKAAAR 839  
Qy 382 AN-----NTTANPRAVDEIKVAKIVLSGLRFGQNTIELSVPPGEGNM-----KV 427  
Db 840 QNLGLNHTITVAQRTALEGOINQATTV-----DGVNTVKTANTALDGMNSLQGAINDKD 894  
Qy 428 APMIGNIYLSNEN 441  
Db 895 ATLKNQNYLDADES 908

RESULT 13  
C89921  
hypothetical protein ebhB [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: C89921  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A69758; MUID:21311952; PMID:11418146  
A:Accession: C89921  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-3890 <KUR>  
A:CROSS-references: GB:BA000018; PID:g13701233; PIDN:BA042528.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: ebhB

Query Match 7.2%; Score 166.5; DB 2; Length 3890;  
Best Local Similarity 19.7%; Pred. No. 0.96;  
Matches 100; Conservative 89; Mismatches 210; Indels 109; Gaps 18;

Qy 17 YGINSSEPT-----QNTSREVSSVOLSEESTFYLCPPVGSSTV 57  
Db 2427 HVSNSPSTLTPRAHTVNTTEIVDYGSNTVAEINNAVQVANKRT-----ATI 2476

Qy 58 -----IRLEFGCMSITKQANPNNGQT-----QLEAARMELTDLINKA 96  
Db 2477 KNGTAMPNINLGGSTTTTPVTYNDGSTBEVQESIFTKADRELITAKNHLDDPVSTEG 2536

Qy 97 MTLASLQDY-----AKIEASLSAYSSEATVNNNLNATLEBOLKMAKTINESAINQANTTKT 152  
Db 2537 KKPGETTQYNNAMHNAQOOINTAKTEAQOVINNEBATPOQVSDALTCTKPAQTKIDQAKA 2596

Qy 153 TPDNEHPN--LVEAYKALKITLEGSRATNLEGSTAYNQIRNNILVDLVN-----KASSLIT 206  
Db 2557 LIQNEEDNSQVTSKNTLQSSVNO-----VPSTRA--GMTQOSIDNTNKKRELETEIT 2647

Qy 207 KTLDLPLNGQTLDSNEITTVNNNINNTLSTINEQKTNADALSNSFIKVIQNNESQFVGT 266  
Db 2648 AAQRVIDNGD-ATAQOISDEKRRVDNALTALNQAHDLADLADTHALEQAVQQLNR---TGT 2703

Qy 267 FTN---ANVQPSNYFVAFASADV-----PV-NYKYARTTWNGDEPSRILA 310  
Db 2704 TTGKKPASTITAYNNSIRALQSDLTSAKUSANAIIQKPIRTVQEVOSALTNNVRVNERLQ 2763

Qy 311 NNSITDVSWISLAGTWTK-----YQFSFNVGPGFNGYLYFKLVKADANNVGLQYKL 366  
Db 2764 AINQVLPLADNSALTKTKLDEBEINKSVTTTGMTQSSIQAYENAKRACQITTTAAQVNI 2823

Qy 367 NNGNVQVEFATSTS-----TSANNTTANPRAVDEIKVAKIVLSGLRFGQNTIELSV 418

Db 2824 NNGDATDOQIAAEKTYVEKYSNLKQALAGLTPDLAFLQTKATQL-----QNDID--QP 2875  
Qy 419 TGEGNMNVKAPMIGNIYLSNENNAKDI 446  
Db 2876 TSTTGMTSASVAAPNDKLSAARTKIOEI 2903

RESULT 14  
AG2422  
hypothetical protein all4935 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: AG2422  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuch  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AG2422  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-661 <KUR>  
A:CROSS-references: GB:BA000019; PIDN:BA076634.1; PID:g17134073; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all4935

Query Match 7.1%; Score 166; DB 2; Length 661;  
Best Local Similarity 22.7%; Pred. No. 0.993;  
Matches 108; Conservative 50; Mismatches 213; Indels 104; Gaps 18;

Qy 1 MHYFRNCIFFLI--VILYGNSS-----PSTQNTSREVSSVOLSEESTFYLCPPV 53  
Db 1 MNLVTQAILGLITTAALGTGLASIKPATVITOTTISAN-VKPIQVAIKEA-----DEAI 54

Qy 54 GSTVIRLEFGCMSITKQANPNNGQTQLEAARMELTDLI-----NAKAMTLASLQDY 105  
Db 55 -----KK---AQEADALKKAEEDLSGIISTRPAKSDANSRLRAQDF 96

Qy 106 AKIEBASLSAYSSEATVNNNLNATLEBOLKMAKTINESAINQANTTKTPDNEHPNLYEAY 165  
Db 97 TOADAGFQIQTALNTATNNRNNAQNALNTATNNRNNAQNALNTATNNRN-----AQ 149

Qy 166 KALKITLEGSRATNLEGSTAYNQIRNNILVDLVNKASSLITKTLDPDLNGTSLDSNEITT 225  
Db 150 NALNTATNNR--NNQNALNTATNN--RNNAQNALNTATNNRNNAQNALN-----TAT 198

Qy 226 VNR--NINNTLSTINEQKTNADALSNSFIKVIQ---NNEQSFVGTFTNANVQPSNYFVA 281  
Db 199 NNRNNAQNALNTATNNNNNAQABVDATRNLAAQRNSQKALQNALNALNQANTRLNTA 258

Qy 282 PSADVTPNVKYKARTYVWNGDEPSRILANTNSITDVSWISLAGTWTKYQFSFNVGPS 341  
Db 259 NTAINTATNNMTANTAFN-----TANNNEST-----ANTAFTATNNFSTA 300

Qy 342 TGYLYFPFKLVKADANNVGLQYKLNNGVQVEFATSTANNTTANPRAVDEIKVAKI 401  
Db 301 -----NTALNTATNNFTALIALDQATKRLNARN-----DPNTANS 337

Qy 402 VLSGLRFGQNTIELSVPTGEGNMNVKAPMIGNIYLSNENNAKDI PGYRRPGTEL 456  
Db 338 NFSRTGMEINATNTNFTNTANTFTTATNTPN--ASSRRTAEDQARQVRETRL 390

RESULT 15  
D90011  
FmbB protease [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: D90011  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

A:Genetic code: SGC3

7.3%; Score 170.5; DB 2; Length 4688;

Query Match Best Local Similarity 21.1%; Pred. No. 0.75; Indels 175; Gaps 25;

Matches 112; Conservative 79; Mismatches 165; Indels 175; Gaps 25;

19 TNSPSTQNTVSRVSVSVQSEESTP-----YICPPVGSYVIREFGCMST 67  
 1387 SNNVSLGNKNTNIEITKRIILVNDQSLISAKIEVDIDNVLNTADKKNIVQLE----- 1440  
 Db 68 TKKDANPN-----NGQTOLEAKMELT-DLINAKMTLASLODYAKIEASISAS- 114  
 1441 NSNNANNMLKLATNAQVAVENKPKFLKPDVLSLKN-----QNYVYKISNSKPTNAYF 1495  
 Db 115 -----AYSEAE-TVNNNLINAT-----LEQKMAKTN 139  
 1496 NPTNKKTNNTIYSDQKIKISLNNINFTSYKPTINKDKSVNIDVDLQVQKOLLANQY 1555  
 Db 140 LESAINQANTDKTT-----PDNEHPNLVEAYKALKITLQEPATNLEGLSTAINQIRNN 193  
 1556 LRLKIKQINDKNTVTDPLIFNN--NAKISFK-LSNLIHNRAYELEGL--YIFDDQNS 1608  
 Db 194 LVDLYNKASLITKTLDPINGSTLLDSNEITTVNNKINNLTSTIN-----EOKTNAD 245  
 1609 VNDMTNNOISFNSKIHKP--KIEFEPSTLTINIDTNNAIKTVSAHNAQVHFKLKTNDE 1664  
 Db 246 ALSNSFIKKVIQNNQSFVGFPTNANVOPSNYSFVAFSADVTPVNYKYARRTWNNGDEPS 305  
 1665 ALENQIVAV-----FAPTNLNDQKV-----E 1689  
 Db 306 SRILANTNSITDVSNTIYSLAG--TWTKYQF--SFSNTGSPGTCYVYFVKLVKADANNV 360  
 1690 AKINNVNTSFGNGBELFENLSGLREBYRLIKVTEKKN-----PKAYELINKNGV 1740  
 Db 361 GLOYKLNNGNVOOVEFAT-----STANNTTANPTPAVDEI-----KVAKIV 402  
 1741 IREYK-NCS-QAYETTKQKFEHXYIDVSSYSTMTQOEIVKIDGIRANNNKKLELV 1797  
 Db 403 LSGLRFGQNTIELSVPTGEGNNKVAEMIGNTIYLSNENNAKTIQYRRPG 453  
 1798 YESNIIQDEIKITV--DNNNN-----SVHLSPDKKEVNLVNLNPKG 1838  
 Db  
 RESULT 11  
 JCG009  
 surface-located membrane protein Imp3 precursor - Mycoplasma hominis  
 C:Species: Mycoplasma hominis  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: JCG009  
 R:Adeloped, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.  
 J. Bacteriol. 178, 2775-2784, 1996  
 A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis Imp gene system  
 A:Reference number: JCG009; MUID:96213016; PMID:8631664  
 A:Accession: JCG009  
 A:Molecule type: DNA  
 A:Residues: 1-1302 <LAD>  
 A:Cross-references: EMBL:X55601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336  
 C:Genetics:  
 A:Gene: Imp3  
 C:Genetic code: SGC3  
 C:Superfamily: surface-located membrane protein Imp3; tetratricopeptide repeat homology  
 C:Keywords: duplication; membrane protein  
 F:1/24/Domain: signal sequence #status predicted <SIG>  
 F:25-1302/Product: surface-located membrane protein Imp3 #status predicted <MAT>  
 F:957-992/Domain: tetratricopeptide repeat homology <TT1>  
 F:993-1026/Domain: tetratricopeptide repeat homology <TT2>  
 F:1089-1120/Domain: tetratricopeptide repeat homology <TT3>  
 F:1154-1199/Domain: tetratricopeptide repeat homology <TT4>

Query Match 7.3%; Score 170; DB 1; Length 1302;  
 Best Local Similarity 22.6%; Pred. No. 0.14; Indels 88; Gaps 18;  
 Matches 93; Conservative 68; Mismatches 163; Indels 88; Gaps 18;

QY 67 ITRKDANPNQ-----TOLEARMELTDLINAKMTLASLODYAKIEASISAYSEATV 122  
 864 ITRKETEPKDKOVKFELEQTRKIDEFIN-----TNKTNPYSTLISELSKRSKNSI 919  
 Db 123 NNNLNATLEQLMKMTNLESAINQANTDKTTFDN-----EHPN-----LVEAYKAL 168  
 920 TNSNNS--DIETATTELQALAKANTDKQADNLTARSTKEQINKSISANTLAKITDK 977  
 Db 169 KTTLEQATNLEGLSTAINQI--RNNVDLVNKASL-----ITKTDPLNGSTLLDSN 221  
 978 DNTIQAKTELEKEVQKANAQVAVSNNTASQASSLDAKTEITIKLETFKDKDVKER 1037  
 QY 222 EITTVNRNINTLSTINEQTNADALNSFIKKVIQNNQSFVGFPTNANVOPSNYSFVA 281  
 1038 ELEQTRKIDEFINT--NKNPNYSTLISELSK--RDSKNSITNSNKSDIETANTEL-- 1092  
 Db 282 FSADVTPVNYKYARRTWNNGDEPSSRIANTNSITDVSNTIYSLAGNTTYQSPFNSGYS 341  
 1093 -----KQALAKAN--TDKAQADNLA-RSTKEQINKSISAN 1125  
 Db 342 TGYLYFPYKLVKADANNVGLQYKLNNGNVOOVEFATSTANNTTANPTPAVDEI--KYA 399  
 1126 T-----LAKITDKONTIQAKTE--LEKEVQKANAQVAVSNNTASQASSLDTKYT 1176  
 Db 400 KIVLSGLRFGQNTIELSVPTGEGNNKVAEMIGNT--YLSNENNAKTIQY 449  
 1177 EITKLETFPKDK-----EAKFNELEKTRGOIQIEFTNTNKN--PNY 1216  
 Db

RESULT 12  
 B89921  
 hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: B89921  
 R:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani, O.; Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.; I. C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: B89921  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-6713 <KUR>  
 A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GND0149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: ebha

Query Match 7.3%; Score 170; DB 2; Length 6713;  
 Best Local Similarity 20.9%; Pred. No. 1.3; Indels 122; Gaps 19;  
 Matches 103; Conservative 77; Mismatches 192; Indels 122; Gaps 19;  
 20 NSSPSTQNTVSRVSVSVQSEESTPYLCPVPVGSYVIREFGCMSTIKKDANPNNGQT 79  
 465 NNLTISINNAQRDLTTKI--DQAT-----TVAGVEANSNG-T 499  
 Db 80 QLEAARMELTDLINAKMTLASLODYAKIEASISAYSEAT-----VNNNLNATLEQK 134  
 500 QLNTRAMNLQNGINDKANTLAS--ENYHADDEKKTATQAVTNAENILNKSGSNLD--- 555  
 Db 135 MAKTNLESAINQANTDKTTFPNEHPNLVEAYKALKITLQEPATNLEGLSTAYNQIRNL 194  
 556 --KAVERNALSQVNRKAKALNGN--NEQASNNATTT--NGLOHLLTQKDKLQOV 608  
 QY 195 VDLVYKAS-SLITITLDPING--GTILDSNEITTVNRNINTLSTINEQKT----- 242  
 609 QQAQVAVDVTVKASANTLNGMGLTIRSIDQNTATNGQVYLATENKNNYNNAVDSA 668  
 Db 243 -----NADALNSFIKKVIQNNQSFVGFPTNANVOPSNYSFVAFSADVTPVNYK 292  
 669 NGVIVATSNPNMDANAIQIATQVSTGNALDGHNLTOAKOT-----ATNAIDGATNLN 723

## RESULT 7

VGBERR

glycoprotein B precursor - Marek's disease virus (strain RA1B)

C:Species: Marek's disease virus

C&gt;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jun-2000

C:Accession: A32402; B32402

R:Rose, L.J.N.; Sanderson, M.; Scott, S.D.; Bime, M.M.; Doel, T.; Milne, B.

J. Gen. Virol. 70, 1789-1804, 1989

A:Title: Nucleotide sequence and characterization of the Marek's disease virus homolog

A:Reference number: A32402; PMID:89293086; PMID:2544666

A:Accession: A32402

A:Molecule type: DNA

A:Residues: 1-865 &lt;ROS&gt;

A:Cross-references: GB:DJ3713; EMBL:D00506; NID:9221836; PDB:BA02866.1; PID:9221837

A:Accession: B32402

A:Molecule type: protein

A:Residues: 250-271,304-330 &lt;ROS2&gt;

C:Superfamily: herpesvirus glycoprotein B

F:1-21/Domain: signal sequence #status predicted &lt;STG&gt;

F:22-865/Product: glycoprotein B #status predicted &lt;GPB&gt;

F:709-728/Domain: transmembrane #status predicted &lt;TM1&gt;

F:732-752/Domain: transmembrane #status predicted &lt;TM2&gt;

F:27,184,332,364,406,425,631/Binding site: carbohydrate (Aan) (covalent) #status predict

Query Match

Best Local Similarity 13.5%; Score 314; DB 1; Length 865;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHYFRNCIFFLIVLILYGTNSPSTQNTYREVSVQLESEESTFYLCPPVSGSTVIRL 60

Db 1 MHYFRNCIFFLIVLILYGTNSPSTQNTYREVSVQLESEESTFYLCPPVSGSTVIRL 60

Qy 61 E 61

Db 61 E 61

RESULT 8

major surface protein (clone pmga1.6) - Mycoplasma gallisepticum (fragment)

C:Species: Mycoplasma gallisepticum

C&gt;Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S51560; S48757

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FBS Lett. 352, 347-352, 1994

A:Title: The organization of the multigene family which encodes the major cell surface P

A:Reference number: S48751; PMID:95010739; PMID:7925999

A:Accession: S51560

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-320 &lt;MAR&gt;

A:Cross-references: EMBL:L28424; NID:9535687; PDB:AAA62415.1; PID:9535688.

C:Genetics:

A:Genetic code: SGC3

Query Match

Best Local Similarity 8.2%; Score 189.5; DB 2; Length 320;

Matches 46; Conservative 11; Mismatches 26; Indels 11; Gaps 3;

Qy 354 AADANNVGLQYKLNNGNVQVEFATSTANNTTANPTPAVDKIKAKIVLGLRFGONTI 413

Db 41 SADSNN-----PTNQSNSQNOAAPASA---MNETPTVDGIVAKVTLTDLKFGSNTI 90

Qy 414 ELSVP-TGEGNNKVAPEMIGNTIYLSSENENADKI 446

Db 91 ELSVPTTDEGTSKVAPEMIGNTIYLSSENENADKI 124

RESULT 9

S41539

fibrinogen-binding protein - Staphylococcus aureus

N:Alternate names: clumping factor

C:Species: Staphylococcus aureus

C&gt;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 15-Oct-1999

C:Accession: S41539; S36630

R:McDevitt, D.; Francoise, P.; Vaudaux, P.; Foster, T.J.

Mol. Microbiol. 11, 237-248, 1994

A:Title: Molecular characterization of the clumping factor (fibrinogen receptor) of Staph

A:Reference number: S41539; PMID:94224142; PMID:8170386

A:Accession: S41539

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-933 &lt;MCD&gt;

A:Cross-references: EMBL:Z18852; NID:9397525; PDB:CAA79304.1; PID:9397526

Query Match

Best Local Similarity 7.6%; Score 176.5; DB 2; Length 933;

Matches 112; Conservative 66; Mismatches 184; Indels 131; Gaps 23;

Qy 2 HYFRNCIFFLIVLILYGTNSPSTQNTYREVSVQLESEESTFYLCPPVSGSTVIRL 52

Db 9 HAIRKSIQVAVLVGLIFGLSLKEADASENSVTSASNSKSSNSVSAAPKT 68

Qy 53 VGSTVIRLEFGCMSITKQANPNNGOTOL--EAPMELTDLINAKAMTLASLDQYAKIEA 110

Db 69 DDTNV-----SDTKSSNVNNGFSTVAQNPQGETTQSSSTNATT----- 108

Qy 111 SLSSAYSEAEVNNNNLNLATLEQLKMTNLESANQANTDKTTPDNHNPVYKALKT 170

Db 109 EETVTEGATTTTNTQANTPATTSNTNABELNQ--TSNETFND--TIVSSVNS-- 162

Qy 171 TLEQRATNLEGLST-----AVNQIRNNLVDLNLK-----ASSLIT 206

Db 163 --PQNSTNAEVSSTQSTTEATPSNNESAPQSDANQVAVNTSAPRMAAFLLA 220

Qy 207 KTLDPNLNGTLLDSNEITTVNRNINNTLSTINEQ---KTN-ADALSNSPIK---KVIO 257

Db 221 VAADAPAGTDI--TNQLTNVGVGIDSGTIVPHQAGVYKLVNGFVSFNSAVKQDPFRITV 279

Qy 258 NNEQSPVGTPTNNAVQSNVSVFAFSADVPVNVKARVTVNDESSHLANTNSITD 317

Db 280 PKEINLVGVSTAKVP-----IMAGDQ---VLANGVIDSD 312

Qy 318 VSMIYSLAG--TNTKYQSFNSVGPSTGYLPYKLVKADANNGLQKLVNGVQVEF 376

Db 313 GNVYITFTDYNTQDDVKATITMFA--YI-----DPENV-----KKTGV----- 350

Qy 377 ATSTANNTTANPTPAVDKIKAKI-----VLGLRFGONTIELSVPTGEGNNKVAPE 429

Db 351 TLATIGISTPANKTVLVDYERKGFYVMSIKGITDQIDKTNNTYRQTIYVNPSSGDNIAP 410

Qy 430 MI--GNIIYLSSEN 441

Db 411 VLTNLKPNTDSN 423

RESULT 10

P82885

hypothetical protein U0482 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C&gt;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: F82885

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.

submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: A82870

A:Accession: F82885

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-468 &lt;GIA&gt;

A:Cross-references: GB:A0002145; GB:AF222894; NID:96899476; PDB:AA30894.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: U0482

Query Match 31.9%; Score 741.5; DB 2; Length 649;  
 Best Local Similarity 40.0%; Pred. No. 3.7e-33;  
 Matches 176; Conservative 65; Mismatches 126; Indels 63; Gaps 12;

```

Oy 53 VGSFVIRLEFGCWSITTK-----DANPNNGQ-----TGLEAARMELTD 90
Db 15 IGSFVMLAAASCTPTPTPNPNPPSGMNGDPTNPGDGMNNAASQELAAARMGLTT 74
Oy 91 LINAKAMTLASLDYAKTEASLSAYSEAEVNNNNLNTLEQLKAKTNLESAINQANTD 150
Db 75 IFDSAKKMLGLVYDKTKQNTLTAKYDAKTVLDNSSSTTONLEAKTRLETAIRTAATS 134
Oy 151 KTFPNEHPNLEAYAKKLTLEORATNLEGLSTAVNQIRNNI VDLYNKASSLITKLD 210
Db 135 KOTPEQHAELVYKKEKLTLSNETATLAPADQVAGIQHMLSGLYDAKAKATTTKLE 194
Oy 211 PLNGSTLDSNEITTVNNRINNTL--STINEOKTNADLSNSFKIVQNNQESFVGFT 268
Db 195 PVEGDP-LTASAVMMANKIYEAIKDEVANPOKENATKLAOSFYKQVYVKEKITGVEBAH 253
Oy 269 NANNVQPSNYSFVAFSADVTVP-----NYKARRTVM-NGDEPSSRIILANT----- 312
Db 254 N-KAQPANYSFVGYSVDITGTTGOTSI PMMDYAKRTITFNSDEP--RSISNTPADQGTM 310
Oy 313 -NSITDVSWIYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVKADANNVGLQYKLNNGV 371
Db 311 AQPILSNVSWIYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVKADANNVGLQYKLNND 367
Oy 372 QQVEPFTSTSA-----NNTTANPTPAVDEIKYAKIVLSGLRFGONTIELSVPTGEGNMK 426
Db 368 -----ATKPSAITFGSDQTMNGKPTVNDIVNAKVTIANLNGSKIEFSVPA-----EK 417
Oy 427 VAPMIGNIYLSNENNAADKI 446
Db 418 VSPMIGNMVLSSSPNNMNKI 437

```

## RESULT 5

A49218  
 hemagglutinin homolog pmGA1.2 - Mycoplasma gallisepticum  
 C:Species: Mycoplasma gallisepticum  
 C>Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 07-Dec-1999  
 C:Accession: A49218  
 R:Markham, P.F.; Glew, M.D.; Whitehead, K.G.; Walker, I.D.  
 Infect. Immun. 61, 903-909, 1993  
 A:Title: Molecular cloning of a member of the gene family that encodes pmGA, a hemagglutinin  
 A:Reference number: A49218; MUID:93162830; PMID:8432610  
 A:Accession: A49218  
 A:Status: Preliminary  
 A:Molecule type: DNA; protein  
 A:Residues: 1-647 <MAR>  
 A:Cross-references: GB:S55216; NID:g265625; PIDN:AA825397.1; PID:g265626  
 A:Experimental source: S6  
 A:Note: sequence extracted from NCBI backbone (NCBIN:125182, NCBI:P:125183)  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match 30.6%; Score 711.5; DB 2; Length 647;  
 Best Local Similarity 39.0%; Pred. No. 1.6e-31;  
 Matches 173; Conservative 63; Mismatches 135; Indels 73; Gaps 13;

```

Oy 53 VGSFVIRLEFGCWSITTK-----DANPNNGQ-----TGLEAARMELTD 90
Db 15 IGSFVMLAAASCTPTPTPNPNPPSGMNGDPTNPGDGMNNAASQELAAARMGLTT 74
Oy 91 LINAKAMTLASLDYAKTEASLSAYSEAEVNNNNLNTLEQLKAKTNLESAINQANTD 150
Db 75 IFDSAKKMLGLVYDKTKQNTLTAKYDAKTVLDNSSSTTONLEAKTRLETAIRTAATS 134
Oy 151 KTFPNEHPNLEAYAKKLTLEORATNLEGLSTAVNQIRNNI VDLYNKASSLITKLD 210
Db 135 KOTPEQHAELVYKKEKLTLSNETATLAPADQVAGIQHMLSGLYDAKAKATTTKLE 194

```

```

Oy 211 PLNGSTLDSNEITTVNNRINNTL--STINEOKTN--ADALSNSFKIVQNNQESFV 264
Db 195 PVEGDP-LTASAVMMANKIYEAIKDEVANPOKENATKLAOSLSISYKKITGVEB----- 249
Oy 265 GFTNANVQPSNYSFVAFSADVTVP-----NYKARRTVM-NGDEPSSRIILANT----- 312
Db 250 ---ANKAQPANYSFVGYSVDITGTTGOTSI PMMDYAKRTITFNSDEP--RSISNTPAD 304
Oy 313 -NSITDVSWIYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVKADANNVGLQYKLN 367
Db 305 GQTMAPLSNWSIYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVNISDQYKLGLEYKLN 364
Oy 368 NGVQOVEPATSTSA-----NNTTANPTPAVDEIKYAKIVLSGLRFGONTIELSVPTGEG 422
Db 365 D-----ATKPSAITFGSDQTMNGKPTVNDIVNAKVTIANLNGSKIEFSVPA--- 413
Oy 423 NNNKVAPMIGNIYLSNENNAADKI 446
Db 414 ---EKVSPMIGNMVLSSSPNNMNKI 435

```

## RESULT 6

S4875  
 major surface protein (clone pmGA1.5) precursor - Mycoplasma gallisepticum (fragment)  
 C:Species: Mycoplasma gallisepticum  
 C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999  
 C:Accession: S4875  
 R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; v  
 FEBS Lett. 352, 347-352, 1994  
 A:Title: The organization of the multigene family which encodes the major cell surface p  
 A:Reference number: S48751; MUID:95010739; PMID:7925599  
 A:Accession: S48755  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-386 <MAR>  
 A:Cross-references: EMBL:128424; NID:g535687; PIDN:AA62419.1; PID:g535692  
 A:Note: the sequence of residues 385-386 and the corresponding nucleotide sequence are nc  
 A:Genetics:  
 A:Genetic code: SGC3  
 A:start codon: GTG

Query Match 24.9%; Score 579; DB 2; Length 386;  
 Best Local Similarity 38.4%; Pred. No. 1.3e-24;  
 Matches 144; Conservative 53; Mismatches 122; Indels 56; Gaps 11;

```

Oy 53 VGSFVIRLEFGCWSIT-----KQDANPN----- 75
Db 15 IGSFVMLAAASCTPTPTPNHEPKPNAPKPPKPNPGGMNGMNGNTNPGGGMND 74
Oy 76 NGOTGLEAARMELTDLINAKAMTLASLDYAKTEASLSAYSEAEVNNNNLNTLEQLK 135
Db 75 NSAQOLAAAKKESLDLATQNSLSTYADAKIKNDLTAAITTAFTASQQAATLEQVK 134
Oy 136 AKTNLESAINQANTDKTFPNEHPNLEAYAKKLTLEORATNLEGLSTRA-YQIIRNNL 194
Db 135 AASTLOTNINAVNAEKKVFENNSSELVAYYNLTKEGENTVTLAANDSANGIGITHL 194
Oy 195 VDLYNKASSLITKLDPLN-AGTLLDSNEITTVNNRINNTL--STINEOKTNADLSNSF 251
Db 195 LSLVNOAKTITTTSTL--LNDAGQSPNKNVVKIKKEITDAINPFLUOQANADMALTSF 252
Oy 252 IKVVIQNNQESFVG---FTNANVQPSNYSFVAFSADVTVP-----VNYKARRTVMNGD 302
Db 253 TKQVL-NDQOLTSSESSEFSMOTQPPGYSFVGYSVDVTTSSNNARPMWNFAQRKWDTN 311
Oy 303 EPPSRILANTNSITDVSWIYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVKADANNV 361
Db 312 RAPLAQTESNNLITDVSWIYSLAGTNTKYQPSFNSYGPSTGYLYFPYKLVQVTD--NVG 369
Oy 362 LQYKLNNGVQVEF 376
Db 370 LQYVLANNTTPKLVNF 384

```



Db 465 NMVLTSSDRDVNK 477

## RESULT 2

S48753

major surface protein (clone pmGAL.3) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48753

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48753; MUID:95010739; PMID:7925999

A:Accession: S48753

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-702 <MAR>

A:Cross-references: EMBL:L28424; NID:9535687; PIDN:AAA62417.1; PID:9535690

C:Genetics:

A:Genetic code: GCG

A:Start codon: GTG

Query Match 32.3%; Score 750; DB 2; Length 702;  
Best Local Similarity 41.4%; Pred. No. 1.4e-33;  
Matches 175; Conservative 62; Mismatches 128; Indels 58; Gaps 11;

QY 73 NNNNGCT---OLEAARHELTDLINAKAMTLASLDYAKIEASLSAYSEAEVNNNNNA 128

Db 69 NNNPNTPEEQOAAARKTLTDLGTEENTNVALVADYAKIOSTLSTAYMTAKTASENTSA 128

QY 129 TLEOLKMAKTNLESAINQANTDKTFEDNEHPMLVEAYALKTTLEORATNLEGSSTAYN 188

Db 129 TLENRASATTLQALIDKANDKRVFDSYNQPLVAINNKTLLSKTTSLEGSSENKYG 188

QY 189 QIRNNLVLDYNKASSLITKTLDPNG--GTLDSNEITVNNINNTLS--TINEQKTNA 244

Db 189 GIKHNLSTLFDGSAITAKTLPTSGERPLEKVN---ANNIGIMAIISPESLKKMKGNA 245

QY 245 DALNSFIKVIQNNQEOSVGTFTNANQPSNYSVASADVTP-----VNYKARRTVW 299

Db 246 DKF-NEFENPLSKERLSTSDTANQEPAMSAVSDLTNSOMLPNNFQORCW 304

QY 300 NCD--EPSRLIANTN-STDVSWIYSLAGTNTKQFSPNYPSTGYLYPYKLVKAAAD 356

Db 305 TSENOQPGKTALVSSPVSATDVSWIYSLAGSGTKTLTFEYIGPDMAFLYLPYKLVKAAAD 364

QY 357 ANNVLQYKLNNGNVQVEF-----ATST-----SAN 383

Db 365 SSSVALQVSLNKTSSKLIFKPAETVSTNTDQSENEVATTSTEARSSYKVLVADEAATS 424

QY 384 NTTNPTPAVDIETKAKIYLSGLRFGQNTIELSVPTGEGNNKVAIPMGINIYLSSENNA 443

Db 425 NNEEMHTPTVSDINIAKVTLSGLTFEGENTIEFSVEG---KVAPMIGMVLTSNSESQ 479

QY 444 DK1 446

Db 480 VK1 482

RESULT 3

S48751

major surface protein (clone pmGAL.1) precursor - Mycoplasma gallisepticum  
N:Alternate names: major hemagglutinin pmGA

C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48751; A44793

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739; PMID:7925999

A:Accession: S48751

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-650 <MAR>

A:Cross-references: GB:U90714; EMBL:L28423; NID:91905869; PIDN:AA50152.1; PID:91905870

R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whitehead, K.G.

Infect. Immun. 60, 3885-3891, 1992

A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepticum

A:Reference number: A44793; MUID:92363591; PMID:1379991

A:Contents: 86

A:Accession: A44793

A:Status: preliminary

A:Molecule type: protein

A:Residues: 26-42 <MA2>

A>Note: sequence extracted from NCBI backbone (NCBI:111017)

C:Genetics:

A:Genetic code: GCG

A:Start codon: GTG

A:Molecule type: DNA

A:Residues: 1-650 <MAR>

A:Cross-references: GB:U90714; EMBL:L28423; NID:91905869; PIDN:AA50152.1; PID:91905870

R:Markham, P.F.; Glew, M.D.; Brandon, M.R.; Walker, I.D.; Whitehead, K.G.

Infect. Immun. 60, 3885-3891, 1992

A:Title: Characterization of a major hemagglutinin protein from Mycoplasma gallisepticum

A:Reference number: A44793; MUID:92363591; PMID:1379991

A:Contents: 86

A:Accession: A44793

A:Status: preliminary

A:Molecule type: protein

A:Residues: 26-42 <MA2>

A>Note: sequence extracted from NCBI backbone (NCBI:111017)

C:Genetics:

A:Genetic code: GCG

A:Start codon: GTG

Query Match 32.1%; Score 745; DB 2; Length 650;  
Best Local Similarity 40.5%; Pred. No. 2.4e-33;  
Matches 177; Conservative 65; Mismatches 139; Indels 56; Gaps 12;

QY 53 VGSTVIRLEFGCHSITKDA-----NPNNGQ-----TOLEAARHELT 89

Db 15 IGSFVMLAASCTTPSPAPNPSPNGNGNINPGGCGMMAAAGELAAARWGLT 74

QY 90 DLINAKAMTLASLDYAKIEASLSAYSEAEVNNNNLNTLEOLKMAKTNLESAINQANT 149

Db 75 TWFDSKAKNLTGLYVDYKKTQNTLTAKYDAKATVDSSSTTONLEAKTLETRIRIAT 134

QY 150 DKTFEDNEHPMLVEAYALKTTLEORATNLEGSSTAYNQRNNLVLDYNKASSLITKTL 209

Db 135 SKQFPDQGHABLVYKVELKTTLSNEFATLAPADAQYAGIKHNLSTGLYDAGKAITTKTL 194

QY 210 DPLNGTGLDSNEITVNNINNTL--STINEQKTNA DALNSFIKVIQNNQEOSVGTFT 267

Db 195 EPVEGDD-LTAGAVTANNTKIVEAIKDEVLPKKNATKLADSFVKVLKEKLTGVEEA 253

QY 268 TNANVQPSNYSPVAFSADVTPV-----NPKYARRTVW-NGDEPSRLIANT----- 312

Db 254 HN-KAOPANYSFPGYSDITGTANGOTSIPNNYAQGITTTNGDEP--RVSNTPVVQC 310

QY 313 --NSITDVSWIYSLAGTNTKQFSPNYPSTGYLYPYKLVKAAADANNVLOKLNNGN 370

Db 311 MAQPLSNVSWIYSLAGTNTKQFSPNYPSTGYLYPYKLVKAAADANNVLOKLNNGN 369

QY 371 VQVVEPFTSANTTN-PTPAVDIETKAKIYLSGLRFGQNTIELSVPTGEGNNKVAIP 429

Db 370 ---TEPSAITFGNEQTNNGKTPVNDINAKVTLANLIFGNNKIEFSVPA-----EKVSP 421

QY 430 MGINIYLSSENNA DK1 446

Db 422 MIGNVLTSSSPNNWVKI 438

RESULT 4

S48752

major surface protein (clone pmGAL.2) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum

C>Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #text\_change 07-Dec-1999

C:Accession: S48752

R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.;

FEBS Lett. 352, 347-352, 1994

A:Title: The organisation of the multigene family which encodes the major cell surface

A:Reference number: S48751; MUID:95010739; PMID:7925999

A:Accession: S48752

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-649 <MAR>

A:Cross-references: EMBL:L28424; NID:9535687; PIDN:AAA62416.1; PID:9535689

C:Genetics:

A:Genetic code: GCG

A:Start codon: GTG

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 10:18:32 ; Search time 42 Seconds  
(without alignments)  
1044.117 Million cell updates/sec

Title: US-09-901-572A-3  
Perfect score: 2324  
Sequence: 1 MHYFRNCIFPLIVLYGTN.....SSNNENADKIPGRPCTFL 456

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	35.6	702	2 S48754	major surface prot
2	750	32.3	702	2 S48753	major surface prot
3	745	32.1	650	2 S48751	major surface prot
4	741.5	31.9	649	2 S48752	hemagglutinin homo
5	711.5	30.6	647	2 A49218	major surface prot
6	579	23.9	386	2 S48755	glycoprotein B pre
7	314	13.5	865	1 VGBRRB	major surface prot
8	189.5	8.2	320	2 S51560	fibronogen-binding
9	176.5	7.6	933	2 S41539	hypothetical prote
10	170.5	7.3	4688	2 F82885	surface-located me
11	170	7.3	1302	1 JC6009	hypothetical prote
12	170	7.3	6713	2 B89921	hypothetical prote
13	166.5	7.2	3890	2 C89921	hypothetical prote
14	166	7.1	661	2 AG2422	hypothetical prote
15	164.5	7.1	2481	2 D90011	FmbB protein (limp
16	162.5	7.0	1237	2 D71850	probable outer mem
17	157.5	6.8	1072	2 A86827	cell surface anti
18	156	6.7	820	2 T17519	hypothetical prote
19	155	6.7	807	2 B71605	hypothetical prote
20	155	6.7	989	2 D89852	1mpl protein - Myc
21	152.5	6.6	1365	2 T30882	toxoid-like outer m
22	152.5	6.6	2399	2 H71879	probable invasin Z
23	151	6.5	2660	2 B85822	subtilisin-like pr
24	150.5	6.5	682	2 S44131	kinesin-related pr
25	148.5	6.4	1073	2 S14032	surface membrane p
26	148	6.4	624	2 PC6003	ser-lys rich hypot
27	147.5	6.3	751	2 T40462	kinesin-like prote
28	147.5	6.3	1085	2 T38378	glycoprotein Vp260
29	147	6.3	1335	2 T17508	

30	146.5	6.3	1051	2 T18351	1mpl protein - Myc
31	144.5	6.2	3194	2 D71917	toxoid-like outer m
32	144	6.2	1107	2 AC0976	probable autotrans
33	143.5	6.2	589	2 B97806	hypothetical prote
34	143	6.2	1487	2 AG2560	hypothetical prote
35	142	6.1	135	2 B49218	hemagglutinin homo
36	142	6.1	719	2 S55119	hypothetical prote
37	142	6.1	868	2 G71691	hypothetical prote
38	141.5	6.1	796	2 T21460	hypothetical prote
39	141.5	6.1	1524	2 S68553	surface layer prot
40	141	6.1	1314	1 TMBIR6	transcription regu
41	139.5	6.0	926	2 AE1130	conserved hypotet
42	139.5	6.0	1116	2 D97001	probable membrane
43	139.5	6.0	1238	2 A64596	hypothetical prote
44	137.5	5.9	568	2 E97066	membrane associate
45	137.5	5.9	1018	2 A32192	fibronectin-bindin

ALIGNMENTS

RESULT 1

S48754 major surface protein (clone pMGAL.4) precursor - Mycoplasma gallisepticum  
C:Species: Mycoplasma gallisepticum  
C:Date: 01-Aug-1995 #sequence\_revision 03-Nov-1995 #ext\_change 07-Dec-1999  
C:Accession: S48754  
R:Markham, P.F.; Glew, M.D.; Sykes, J.E.; Bowden, T.R.; Pollocks, T.D.; Browning, G.F.; V  
FEBS Lett. 352, 347-352, 1994  
A:Title: The organisation of the multigene family which encodes the major cell surface pr  
A:Reference number: S48751, PMID:95010739; PMID:7925999  
A:Accession: S48754  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-702 <MAR>  
A:Cross-references: EMBL:128424  
C:Genetics:  
A:Genetic code: SGC3  
A:Start codon: GTG

Query Match 35.6%; Score 828; DB 2; Length 702;  
Best Local Similarity 40.6%; Pred. No. 7.9e-38;  
Matches 176; Conservative 86; Mismatches 125; Indels 46; Gaps 9;

QY	50	PPVGSIVIRLEFGCSITKQANPNNG-----OTQLEAKMELDTLNAKMTLASIQ	103
DB	54	PNFGGMMGMMGG-----NTNPGCGGTDAQAQAAKKELDLATQNSNLSTVA	106
QY	104	DYAKIEASLSAYSEAEETVNNNNINATLEOLKMAKTNLESAINQANTDXTFEDNEHPLVE	163
DB	107	DYANIQNTLTAAVTTAKSTSDNTSAILQYKASATSTLQTHIDTPAASKISFDEKNPELIK	166
QY	164	AYKAKTTLLEORATNLEGLSTAYNOIRNNLVLDYKFASSLLITKTDPLNGCTLLDSNEI	223
DB	167	AYNKLKTLKMRNLSGLTDSNFATIKTNLTALYQGXIVTKTDLPLM-GTAINNLSAV	225
QY	224	TVVENINNTLSTNEQKTNADALNSFIKVIQNEQSGVGFETNANVQPSVFAFS	283
DB	226	SOANTNISNAVSKELETKVATVATLATSFAKEVLKNTLTGDT-TNNQOQPGNYSFGYS	284
QY	284	ADVTP-----VYKVARRTVMNGD-----EPSSRIILANTNSITDVSMTSLACTNTRYQ	332
DB	285	VDTVTGSGNAPNWSFAQRKKTWTSNTDILSGPOPAEGENQSAVDVMTVLTGMGAKYS	344
QY	333	FSFSNYGPGSTGYLYFPYKLVKADANNVGLQKLNNGNVQOVERATS-----TSAN	383
DB	345	LTENVYGPSTGFLVFPYKLVNNSSDSKVALEYKKNESAVKTIIDSPQTSFVADATREN	404
QY	384	N-----TTANPPAVNVEIVAKIVLSGFRGQNTIELSP-TGEGNMNVAPMIG	432
DB	405	NRSTAAPQSGSTEINPATLDDIKIAKVTLSNLKFGSNTIEFSPTTAKESTSVAPMIG	464
QY	433	NYLSSNNENADK 445	

TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5835

Query Match 6.7%; Score 156; DB 9; Length 2434;  
Best Local Similarity 22.9%; Pred. No. 0.0033;  
Matches 98; Conservative 68; Mismatches 164; Indels 98; Gaps 20;

```

QY 79 TQLEAARELDLIAAKAMTLASLDYAKIEASLSAYSEAEVTANNLNATLEQLKMAKT 138
DB 1003 SELNTPAMSNLQNGINDEAATKHA-QKYTDADREKOTAYNDAVTAATAKTLIDKTAGSNDKKA 1061
QY 1139 NLESAINQANTDKTTFD-----NEHPN-----LVEAYKALKTTLEQRATNLEGL 182
DB 1062 AVEQALQVRNTAKTLNGDERINEAKTRAKQOVATMSHITDAQXANLTSQIESGTTVAGV 1121
QY 163 S-----STAYNQIRNNL-----VDLYN-----KASLITTKTLD 210
DB 1122 QGIGANAGTLDQAMNQLRQSIASKDATKSEEDYQDANADLQAVYNDVATNAEGTISATNN 1161
QY 211 PLNGGTLDSNEITTVNRINNTLTIN--EQKTADALSNSEFIKKVIQ--NNEQSFVGTG 267
DB 1162 P-----EMNPDTINQKASQVNSAKSALNGDEKLAQAQOTAKSDIGRLTDLNNAQR--TA 1233
QY 268 TNANV-OPSNYSFV-----AFSADVTPVNYKYA-----RRTV--WNGDEPSSRIILAN 311
DB 1234 ANAEVDQAPNLAAVYRAKRAKATSLNTAMGNLKHALLAEKDNTRKSVNYTDDAQPKQ--AY 1291
QY 312 TNSITDVSWIYSLAGTN--TRYQFSFSNYGPGTGYLYFPYKLVKADANNVGLQYKLN 368
DB 1292 DTAAVTQAEKITVANGSNNAETQVQAALNQAKNDLNGDNKVAQAKESAKRALASYNL 1351
QY 369 GNVQ-----QVEFATS-----TSANNTANPTPAVDEIKVAKIVUSGLRFGONTIELSVP 418
DB 1352 NNAQSTAAITSQIDNATTVAGVTAQNTANEINTAMGQLQ-----NGIN-DQNTVKKQOVN 1404
QY 419 TSEGNNNK 426
DB 1405 FTDADQK 1412

```

Search completed: August 14, 2003, 10:25:26  
Job time : 29 secs

```

; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13083
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
;
US-09-815-242-13083
Query Match      6.8%; Score 159; DB 9; Length 1048;
Best Local Similarity 23.8%; Pred. No. 0.00055;
Matches 88; Conservative 52; Mismatches 147; Indels 82; Gaps 19;

QY 79 TQLEARMELTDLINKAMTILASLDYAKIEASLSAYSEA-ETVNNINATLEQLKAK 137
DB 85 TDLNTMGNIQGAINDGQTLLS-QNYQATPSKATYAYNAQAKDILINKNGQNK-TK 142
QY 138 TNLSEAINQANTDKTTFDEHEPMLVEAYALKTLLEQRAITNLEGLSTAYNOIRNNVLDL 197
DB 143 DQYTEAMNQVNSAKNNLDG---TRLDOAKQTAKQCLNNMTHLTQAKTMLTQO--- 193
QY 198 YNKASSLITKTLPLNGCTILDSNEITTYNRN-INNTLSTINEQKNADA--LSNSFI-- 252
DB 194 -----INSGLTVAGVQTVQSANLTDQAMTTLKQSLANKQATASSEDYDA 239
QY 253 ---KKVIONNEQSFVGTFTNANYQPS-NYSFVAFSADVTVPVYKARRTVMNGDE----- 303
DB 240 NNDQOTRYNNAVAALERTIYANSPENPSTIQKAE-QVN---SKTALNGDEMLTA 294
QY 304 -PSSRIANT-NSITD-----VSWIYS---LACTNTKIQPSNSNGPSTGYLYFPYKLV 352
DB 295 KQNAKTYLNTLTSITDAQKNLLISQITSATRVSGVDITVKQNA----- 336
QY 353 KAADANNVGYQYKLNNNGVOVEPATSTANNTTANPTPAVDE-IVAKIVL---SGLRF 408
DB 337 QHLDQAMASLONGINN---ESQVKSSEKTRDADITKQOEYDNAITRAKAILINKSGPMT 392
QY 409 GQNTIELSV 417
DB 393 AQNAVEAAL 401

RESULT 14
US-09-820-843A-108
; Sequence 108, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEI
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 807
; TYPE: PRT
```

```

; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: hypothetical protein
; NAME/KEY: misc.feature
; OTHER INFORMATION: gi|3845292
;
US-09-820-843A-108
Query Match      6.7%; Score 156; DB 11; Length 807;
Best Local Similarity 23.8%; Pred. No. 0.00066;
Matches 73; Conservative 48; Mismatches 124; Indels 62; Gaps 14;

QY 122 VNNINATLEQLKMAKT-----NLSEAINQA--NTDK-----TFEDNEHPNLVEAYALK 170
DB 468 INNINISYT--QKKNNNINFINHESPINQCHNNTEFKVNNPDIFNEPTNKQKKKEK 525
QY 171 TLEQATNLEGLSTAYNOIRNNVLDLYNKASSLITKTLPLNGTLLDSNEITTVNRI 230
DB 526 NIHFNNNNNNNNKKCLYKDINQ---DHNSLIINTNQFDDH-----NNVKTEDWL 574
QY 231 ---NNTLSTINEQKNADALNSFIKVIQNNNEQSFVGTFTNANYQPSNYSFVAFSADVT 287
DB 575 QKHNNPMSQVSKQNNKNNKNSHLKKQJNIN-----TNNNDKNNSHISGVAYD 626
QY 288 PVNYK-----YARTVWNGDEPSS---RLANTNSITDVSWIYSLAGTNTYQ-FSPSN 337
DB 627 DNLLKSHADNSMEIYTKGKKKTKTKKKKINNINSVNNNNINNNINNNIISMN 686
QY 338 YPSTGYLYFPYLVAAANNVGLQY-----LNNGNVQVFPATSTSAN 383
DB 687 VNNMNNMYPFPVNVIO-KQDSNIALLYNNKPNIDFNNFQLNHINNMIONNIMTNVMLN 745
QY 384 N--TTAN 388
DB 746 NNLTTSN 752

RESULT 15
US-09-815-242-5835
; Sequence 5835, Application US/09815242
; Patent No. US20020061699A1
; GENERAL INFORMATION:
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5835
; LENGTH: 2434
```

PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 12996  
LENGTH: 6281  
TYPE: PRF  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12996

Query Match 7.3%; Score 169.5; DB 9; Length 6281;  
Best Local Similarity 21.9%; Pred. No. 0.0011;

Matches 96; Conservative 67; Mismatches 179; Indels 97; Gaps 16;

75 NNGOTLEARMELTDLINAKAMTLASLDYAKIEASISAYSEAET-----VNNNLNAT 129  
266 SNTSQTQMTAMANTLNGINDKNTILAS-ENTHDADSDKKTATQATNAENILNNGSSN 324

130 LEOLMAKTNESAINQANTDKTFDNEHPNIVEAYKALKTTLEORATNLSGSTAVNQ 189  
325 LD-----KTAVERNALSQVANAKGALNGNH-NLEQAKSNANTTI-----NGLOHLLTAQKDX 374

190 IRNNLVLDLYNKS-SLITFTLDPLNG--CTLLDSSEITTVNRNINNTLSTINEQCT----- 242  
375 LKQVQQAQNAVGVDTFVSSANTLNGAMGTLSNLSIQDNATKQNYLDATERNKTNNYNN 434

243 -----NADALSNSFTKQVIONNEQSFVGFPTTANVQPSNYSFAFSADVT 287  
435 AVDSANGVYNATSNRMDANAINOJAIVTSTKMLDGHNLTOAKQT-----AINALDQ 489

288 PVNYKARRTVNGDEPSSRIANTNSITDVSIMYSLA-----GNTTRYQSF 335  
490 ATNLNKAQKDALKAQVTSARVANVTSIQOTANEILNTAMGQLQHGIDDENATKOTOKYRD 549

336 SNYSGSTGLYFPYKLYKADA-----NNVGLQYKN-----NGNVQVYEF 376  
550 AEOSSKRTAY-----DQVAALAKAILNKQGTGNSDKAAVDBALQOVSTYKDALNGDAKLA 605

377 ATSTSAN-----NTTANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGEGNAN----- 425  
606 KAAAKQNLGTNLHITNAQRTDLEGOINQATVY-----DQVNTVKTNANTLDGANNLSQGS 660

426 ---KVAPMIGNIYLSNEN 441  
661 INDKDATTLENNYLDADDS 679

RESULT 12  
US-09-815-242-5885  
Sequence 5885, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011a

CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 5885  
LENGTH: 1029  
TYPE: PRF  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5885

Query Match 6.8%; Score 159; DB 9; Length 1029;  
Best Local Similarity 23.8%; Pred. No. 0.00054;

Matches 88; Conservative 52; Mismatches 147; Indels 82; Gaps 19;

79 TOLEARMELTDLINAKAMTLASLDYAKIEASISAYSEA-ETVNNNLNATLEOLMAK 137  
73 TDLNTAMGNLQGAINDQETILNS-QNYQDATPSKRTATVAVQAKDILKNSGQNK-TK 130

138 TNLSEAINQANTDKTFDNEHPNIVEAYKALKTTLEORATNLSGSTAVNQIRNNLVLD 197  
131 DQVEANQVNSAQNNDG-----TRLLDQAKOTAKQOLNMHTLITTAQNTLNTNQ----- 181

198 YNKASLITFTLDPLNGTLDSNEITTVNRN-INNTLSTINEQKTNADA--LSNSFT-- 252  
182 -----INGTIVAGQVQVQSNANTLDAQMNTLRQSLANKDKATKASSEDVDA 227

253 ---KKVIONNEQSFVGFPTTANVQPS-NYSFVAFSADVTPVNYKARRTVNGDE----- 303  
228 NNDQRTAYNNAVAALFTIINANSNPEMNPSTIQKAE--QVN--SEKTLNGDENLTA 282

304 -PSSRIANT-NSITD-----VSMYIS---LQGTNKKQPSFNSPSTGYLYFPYKLV 352  
283 KQNKTYLNTLTSITDQKNNLISQITSATRVSGVDVTKONA----- 324

353 KAADANNVGLQYKLNQVQVEFATSTANNTTANPTPAVDE-IKVAKIVL---SGLRF 408  
325 QHLDQAMASLQNGINN-----ESQVKSSEKTRDADTNKQOEVDNATIAAKAILNKSTGPT 380

409 QNTTIELSV 417  
381 AQNAVEAL 389

RESULT 13  
US-09-815-242-13083  
Sequence 13083, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011a

TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5639

Query Match 7.3%; Score 170; DB 9; Length 2086;

Best Local Similarity 20.7%; Pred. No. 0.0002;  
Matches 103; Conservative 77; Mismatches 195; Indels 122; Gaps 19;

```

QY 17 YGTTSSPSTQNTSVSEVSVQLSSEESTFYLCPPVGSVIRLEFGCMSTTKKDANPN 76
DB 886 HALLNLTSTINNKKRDLTKI---DQAT-----TVAGVEAVSNT 921
QY 77 GGTGLEAAMELTDLINAKAMTLASLDYAKIEASLSAYSEAT-----VNNNLATLE 131
DB 922 G-TQNTAMANIQNGINDKANTLAS-ENYHDADSDKKTAYTQAVTNAMENILNKSGSNLD 979
QY 132 QLMKAKTNLESAINQANDKTTFDNEHENVLYEAYKALKTTLEORATNLEGLSTAYNQIR 191
DB 980 ---KAAVENMLSQVTNAKALNGNH-NLEQAKSNANTTI---NGLOHLTTAQKDK 1029
QY 192 NNLVLDLYNKAS-SLITKTLDPUNG--GTLDSNEITTVNRNINNTLSTINEQKT----- 242
DB 1030 QQVQAQAVAGVDVTSSANTLNGAMGTLRNSIQNTATNNQGYLDATESNKTNNYNAV 1089
QY 243 -----NADALSNSFIKAVIQNNEQSFVGTFTNANVQPSNYSFVAFSADVTPV 289
DB 1090 DSANGVYNATSNPNMDANAINQIATQVSTKVALDGTNLTQAKQT-----ATNALDQAT 1144
QY 290 NYKVARRTVMNGDEPSRIANTNSITDVSMTYSLA-----GTNTRYQSPFSN 337
DB 1145 NNNKQKDALKAQVTSAGRVANVTSIQQTANIELNTMGOLOHGIDENAKOTOKYRDAE 1204
QY 338 YGPGTGYLPFYKLYKADA-----NNVGLQYKLN-----NGVQOVEPAT 378
DB 1205 QSKKTAY---DQAVAAKAILNKQTSNSDKAAVDRALQGYVSTKDALNGDAKLAEAKA 1260
QY 379 STSAN-----NTNANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGSGNN----- 425
DB 1261 AAKQNLGTLNHTNAQRTALEQINQATTV---DGVNTVKTNANTLDGAMNSIQGSIN 1315
QY 426 -KVAPMIGNIYLSNEN 441
DB 1316 DKDATTNRNQYLDADDES 1332

```

RESULT 10  
US-09-815-242-12610  
Sequence 12610, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyckind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/220,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12610  
LENGTH: 5795  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12610

Query Match 7.3%; Score 170; DB 9; Length 5795;  
Best Local Similarity 20.7%; Pred. No. 0.00091;  
Matches 103; Conservative 77; Mismatches 195; Indels 122; Gaps 19;

```

QY 17 YGTTSSPSTQNTSVSEVSVQLSSEESTFYLCPPVGSVIRLEFGCMSTTKKDANPN 76
DB 3163 HALLNLTSTINNKKRDLTKI---DQAT-----TVAGVEAVSNT 3198
QY 77 GGTGLEAAMELTDLINAKAMTLASLDYAKIEASLSAYSEAT-----VNNNLATLE 131
DB 3199 G-TQNTAMANIQNGINDKANTLAS-ENYHDADSDKKTAYTQAVTNAMENILNKSGSNLD 3256
QY 132 QLMKAKTNLESAINQANDKTTFDNEHENVLYEAYKALKTTLEORATNLEGLSTAYNQIR 191
DB 3257 ---KAAVENMLSQVTNAKALNGNH-NLEQAKSNANTTI---NGLOHLTTAQKDK 3306
QY 192 NNLVLDLYNKAS-SLITKTLDPUNG--GTLDSNEITTVNRNINNTLSTINEQKT----- 242
DB 3307 QQVQAQAVAGVDVTSSANTLNGAMGTLRNSIQNTATNNQGYLDATESNKTNNYNAV 3366
QY 243 -----NADALSNSFIKAVIQNNEQSFVGTFTNANVQPSNYSFVAFSADVTPV 289
DB 3367 DSANGVYNATSNPNMDANAINQIATQVSTKVALDGTNLTQAKQT-----ATNALDQAT 3421
QY 290 NYKVARRTVMNGDEPSRIANTNSITDVSMTYSLA-----GTNTRYQSPFSN 337
DB 3422 NNNKQKDALKAQVTSAGRVANVTSIQQTANIELNTMGOLOHGIDENAKOTOKYRDAE 3481
QY 338 YGPGTGYLPFYKLYKADA-----NNVGLQYKLN-----NGVQOVEPAT 378
DB 3482 QSKKTAY---DQAVAAKAILNKQTSNSDKAAVDRALQGYVSTKDALNGDAKLAEAKA 3537
QY 379 STSAN-----NTNANPTPAVDEIKVAKIVLSGLRFGONTIELSVPTGSGNN----- 425
DB 3538 AAKQNLGTLNHTNAQRTALEQINQATTV---DGVNTVKTNANTLDGAMNSIQGSIN 3592
QY 426 -KVAPMIGNIYLSNEN 441
DB 3593 DKDATTNRNQYLDADDES 3609

```

RESULT 11  
US-09-815-242-12996  
Sequence 12996, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyckind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21

US-08-781-986A-5249

Query Match 7.6%; Score 176.5; DB 8; Length 936;

Best Local Similarity 22.5%; Pred. No. 1.9e-05;

Matches 111; Conservative 64; Mismatches 187; Indels 131; Gaps 22;

```
OY 2 HYFRANCFFLLVILV-----VGTNSS-----PSTONVIREVVSVOLSEESTFYLCPP 52
DB 18 HAIRKSGVAVSLVGTILGFLSSKEADSENSVTOSDSMSNSKSDSSVSAAKT 77
OY 53 VGSIVIRLEFGCMSITKKDANPNNGOTOL--EAAEMELTDLINAKAMTLASIDYAKTEA 110
DB 78 DDITNV-----SDTKTSSNTNNGETSVAAQPAQOETTSSTNATL----- 117
OY 111 SLSSVSEAEFTYNNNLNLTLEQLKAKTNLSAIIQANTDKTTPNEHPNLVEAKAKT 170
DB 118 EETPVGTGATTTTNTQANTPATTOSSNTAEELVNGTSNETTSNDT---NTVSSVNS--- 171
OY 171 TLEQRTNLEGLST-----AYNOIRNMLVDLYNK-----ASSLIT 206
DB 172 --PQNSTAENVTQODTSTEATPSNNSAPOSTDASNKDVNOAVNTSAPMRAPSLAA 229
OY 207 KTLDELNGTLLDSNEITTVNNINNTLSTINEQ---KTN-ADALSNSFIK---KYIQ 257
DB 230 VAADAPVAGTDI--TQLTNTVTVGIDSGTIVPHQAGVYKANGFSPNSAVKGDFTKITV 288
OY 258 NNEQSFVGTFTNANQPSNYSFVAFSADVTVPNYKAKARTVNGDEPSSRIIANTNSITD 317
DB 289 PEELNLNGVTSTAKVPP-----IMAGDQ---VLANGVIDSD 321
OY 318 VSMIYSLAG--TWTKYQFSFSNYGSTGYLYFPYKLVKADANNVGLQYKLNNGVQOYEF 376
DB 322 GNVITTFPDYNTKODVAKATLMPA--YI-----DPENV-----KTKGNV----- 359
OY 377 ATSTSANNTANPTPAVDEIKVAKI-----VLSGLRGQNTIELSVPTGEGNNKVP 429
DB 360 TLATIGSTTANKTVLVDEYKGYFNLSIKGTIDIDKTNNTYRQTIYVNSGDNVLAIP 419
OY 430 MI-GNIIYSSNEN 441
DB 420 VLTGNLKPVTDSN 432
```

## RESULT 8

US-10-056-052-2

Sequence 2, Application US/10056052

Publication No. US2003009656A1

GENERAL INFORMATION:

APPLICANT: PATTI, Joseph M

APPLICANT: HUTCHINS, Jeff T

APPLICANT: DOMANSKI, Paul

APPLICANT: PATEL, Pratiksha

APPLICANT: HALL, Andrea

TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO THE CLFA PROTEIN

FILE REFERENCE: P07069US04/BAS

CURRENT APPLICATION NUMBER: US/10/056,052

CURRENT FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: 60/308,116

PRIOR FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: 60/298,413

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: 60/274,611

PRIOR FILING DATE: 2001-03-12

PRIOR APPLICATION NUMBER: 60/264,072

PRIOR FILING DATE: 2001-01-26

NUMBER OF SEQ ID NOS: 20

SOFTWARE: Patencin version 3.1

SEQ ID NO 2

LENGTH: 520

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-10-056-052-2

Query Match 7.4%; Score 173; DB 15; Length 520;

Best Local Similarity 22.7%; Pred. No. 1.5e-05;

Matches 105; Conservative 62; Mismatches 173; Indels 122; Gaps 21;

```
OY 24 STONVIREVVSVOLSEESTFYLCPPVGSIVIRLEFGCMSITKKDANPNNGOTOL-- 81
DB 1 SENSVTOSDSMSNSKSDSSVSAAPKTDITNV-----SDTKTSSNTNNGETSVAAQ 52
OY 82 EAAEMELTDLINAKAMTLASIDYAKTEBASLSSVSEAEFTYNNNLNLTLEQLKAKTNLE 141
DB 53 NPAQOETTSSTNATL-----EETPVGTGATTTTNTQANTPATTOSSNTAEELVNGTS 100
OY 142 SAIIQANTDKTTPNEHPNLVEAKAKTLEQRTNLEGLST-----AYNOIRN 192
DB 101 ELVNG--ISNETTFD--TNTVSSVNS-----PQNSTAENVTQODTSTEATPSNNSAP 152
OY 193 NLVDLYNK-----ASSLITKTLDELNGTLLDSNEITTVNNINNTLSTI 237
DB 153 QSTDASNKDVNOAVNTSAPMRAPSLAAVAADAPVAGTDI--TQLTNTVTVGIDSGTIVY 211
OY 238 NEQ---KTN-ADALSNSFIK---KYIQNNEQSFVGTFTNANQPSNYSFVAFSADVT 288
DB 212 PHQAGVYKANGFSPNSAVKGDFTKITVPEELNLNGVTSTAKVPP----- 257
OY 289 VNYKAKARTVNGDEPSSRIIANTNSITDVSMIYSLAG--TWTKYQFSFSNYGSTGYLYF 347
DB 258 -----IMAGDQ---VLANGVIDSDGNVITTFPDYNTKODVAKATLMPA--YI-- 300
OY 348 PYKLVKADANNVGLQYKLNNGVQOYEFATSTSANNTANPTPAVDEIKVAKI----- 401
DB 301 -----DPENV-----KTKGNV-----TLATIGSTTANKTVLVDEYKGYFNLSIK 342
OY 402 -VLSGLRGQNTIELSVPTGEGNNKVPML-GNIIYSSNEN 441
DB 343 GTIDQIDKTNNTYRQTIYVNSGDNVLAIPVLTGNLKPVTDSN 384
```

## RESULT 9

US-09-815-242-5639

Sequence 5639, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyckind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5639

LENGTH: 2086

SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 23  
 LENGTH: 357  
 TYPE: PRT  
 ORGANISM: Mycoplasma gallisepticum  
 FEATURE: Modified TTT-1 portion (downstream of BglI) of  
 OTHER INFORMATION: pMZ40K-S  
 US-10-131-591A-23

Query Match 76.2%; Score 1770; DB 15; Length 357;  
 Best Local Similarity 98.0%; Pred. No. 3.6e-133;  
 Matches 350; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 89 TDLINAKAMTLASIQDYAKIEASISAYSEAEVTNNNNLNLLEQLAKAKTLVLSAIIQAN 148  
 DB 1 TDLINAKAMTLASIQDYAKIEASISAYSEAEVTNNNNLNLLEQLAKAKTLVLSAIIQAN 60  
 QY 149 TDKTFENDEHPVLVAYVAKLTLEORATNEGLSTAYNOIRNNLVLYNKASLITKT 208  
 DB 61 TDKTFENDEHPVLVAYVAKLTLEORATNEGLSTAYNOIRNNLVLYNKASLITKT 120  
 QY 209 LDPLNGTLLDSNEITTTNNRINNTLSTINEQKTNADALSNSFIKVIQNNQSGVGTFT 268  
 DB 121 LDPLNGTLLDSNEITTTNNRINNTLSTINEQKTNADALSNSFIKVIQNNQSGVGTFT 180  
 QY 269 NANVOPSNVSVAFESADVTPVNYKYARKTWNNGDEPSSRIILANTNSITDVGMIYSLAGTN 328  
 DB 181 NANVOPSNVSVAFESADVTPVNYKYARKTWNNGDEPSSRIILANTNSITDVGMIYSLAGTN 240  
 QY 329 TKYQSFNSVSGTGYLYEPYKLVKAADANNVGLQYKLNNGVQVEPATSTANNTAN 388  
 DB 241 TKYQSFNSVSGTGYLYEPYKLVKAADANNVGLQYKLNNGVQVEPATSTANNTAN 300  
 QY 389 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGMNNKVPAMIGNIYLSNENNDK 445  
 DB 301 PTPAVDEIKVAKIVLSGLRFGQNTIELSVPTGEGMNNKVPAMIGNIYLSNENNDK 357

RESULT 5  
 US-10-131-591A-5  
 Sequence 5, Application US/10131591A  
 Publication No. US20030059799A1  
 GENERAL INFORMATION:  
 APPLICANT: Nippon Zeon Co., Ltd.  
 TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof  
 FILE REFERENCE: J209  
 CURRENT APPLICATION NUMBER: US/10/131,591A  
 CURRENT FILING DATE: 2002-08-15  
 NUMBER OF SEQ ID NOS: 79  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 5  
 LENGTH: 62  
 TYPE: PRT  
 ORGANISM: Marek's disease gammaherpesvirus  
 FEATURE:  
 OTHER INFORMATION: MDVgB signal  
 US-10-131-591A-5

Query Match 13.8%; Score 320; DB 15; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-18;  
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLVILVIGTSSPSTQNTSREVVSSVQSEESTFYLCPPVGSIVIRL 60  
 DB 1 MHYFRNCIFFLVILVIGTSSPSTQNTSREVVSSVQSEESTFYLCPPVGSIVIRL 60  
 QY 61 EF 62  
 DB 61 EF 62

RESULT 6

US-10-131-591A-6  
 Sequence 6, Application US/10131591A  
 Publication No. US20030059799A1  
 GENERAL INFORMATION:  
 APPLICANT: Nippon Zeon Co., Ltd.  
 TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof  
 FILE REFERENCE: J209  
 CURRENT APPLICATION NUMBER: US/10/131,591A  
 CURRENT FILING DATE: 2002-08-15  
 NUMBER OF SEQ ID NOS: 79  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 62  
 TYPE: PRT  
 ORGANISM: Marek's disease gammaherpesvirus  
 FEATURE:  
 OTHER INFORMATION: Modified VgB signal  
 US-10-131-591A-6

Query Match 13.3%; Score 308; DB 15; Length 62;  
 Best Local Similarity 96.8%; Pred. No. 1.1e-17;  
 Matches 60; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MHYFRNCIFFLVILVIGTSSPSTQNTSREVVSSVQSEESTFYLCPPVGSIVIRL 60  
 DB 1 MHYFRNCIFFLVILVIGTSSPSTQNTSREVVSSVQSEESTFYLCPPVGSIVIRL 60  
 QY 61 EF 62  
 DB 61 EF 62

RESULT 7  
 US-08-781-986A-5249  
 Sequence 5249, Application US/08781986A  
 Publication No. US2003005436A1  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5255  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/781,986A  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benson, Bob  
 REGISTRATION NUMBER: 30,446  
 REFERENCE/DOCKET NUMBER: PB248BP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 5249:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 936 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein



```

Db      121  TNNNNLNTLEQLQAKNTLSAIOANTDKTTFNEHNPVLEAVKAKLTILEGRATYLE 180
Qy      181  GLSSTAVNQIRNNLVLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Db      181  GLSSTAVNQIRNNLVLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Qy      241  KTNADALNSFIKKYIIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKARRTVM 300
Db      241  KTNADALNSFIKKYIIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKARRTVM 300
Qy      301  GDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Db      301  GDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Qy      361  GLQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Db      361  GLQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Qy      421  EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGTFL 456
Db      421  EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGTFL 456

```

# RESULT 2 US-09-147-052-4

```

; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A10RU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147, 052
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
; US-09-147-052-4

```

Query Match 86.4%; Score 2008; DB 9; Length 1086;  
Best Local Similarity 99.5%; Pred. No. 1.9e-151;  
Matches 384; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Qy      61  EFGCSITKDKANPNNGQTOLEARMELTDLINAKAMTLASLODYAKIEASLSAYSSEAE 120
Db      691  EFGCSITKDKANPNNGQTOLEARMELTDLINAKAMTLASLODYAKIEASLSAYSSEAE 750
Qy      121  TNNNNLNTLEQLQAKNTLSAIOANTDKTTFNEHNPVLEAVKAKLTILEGRATYLE 180
Db      751  TNNNNLNTLEQLQAKNTLSAIOANTDKTTFNEHNPVLEAVKAKLTILEGRATYLE 810
Qy      181  GLSSTAVNQIRNNLVLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Db      811  GLSSTAVNQIRNNLVLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 870
Qy      241  KTNADALNSFIKKYIIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKARRTVM 300
Db      871  KTNADALNSFIKKYIIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKARRTVM 930
Qy      301  GDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Db      931  GDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 990

```

```

Qy      361  GLQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Db      991  GLQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 1050
Qy      421  EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGTFL 456
Db      1051  EGNMNVKAPMIGNIYLSNENNADKIPGYRRPGTFL 1086

```

# RESULT 3 US-10-131-591A-12

```

; Sequence 12, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131, 591A
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Mycoplasma gallisepticum
; FEATURE:
; OTHER INFORMATION: TTM-1 portion of pNZ40K-S
; US-10-131-591A-12

```

Query Match 83.3%; Score 1936.5; DB 15; Length 384;  
Best Local Similarity 99.7%; Pred. No. 2.1e-146;  
Matches 384; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

Qy      61  EFGCSITKDKANPNNGQTOLEARMELTDLINAKAMTLASLODYAKIEASLSAYSSEAE 120
Db      1  EF-CMSITKDKANPNNGQTOLEARMELTDLINAKAMTLASLODYAKIEASLSAYSSEAE 59
Qy      121  TNNNNLNTLEQLQAKNTLSAIOANTDKTTFNEHNPVLEAVKAKLTILEGRATYLE 180
Db      60  TNNNNLNTLEQLQAKNTLSAIOANTDKTTFNEHNPVLEAVKAKLTILEGRATYLE 119
Qy      181  GLSSTAVNQIRNNLVLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 240
Db      120  GLSSTAVNQIRNNLVLYNKASSLITKTLDPNGGTLDSNEITTVNRNINNTLSTINEQ 179
Qy      241  KTNADALNSFIKKYIIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKARRTVM 300
Db      241  KTNADALNSFIKKYIIONNEQSFVGTFTNANVQPSNYSFVAFSADVTVPVNYKARRTVM 300
Qy      301  GDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 360
Db      240  GDEPSSRIANTNSITDVSWIYSLAGTNTKYQFSFSNYPSTGYLYFPYKLVKADANNV 299
Qy      361  GLQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 420
Db      300  GLQYKLANNGNVQVEFATSTSANNTTANPTPAVDEIKVAKIVLSGLRGQNTIELSVPTG 359
Qy      421  EGNMNVKAPMIGNIYLSNENNADK 445
Db      360  EGNMNVKAPMIGNIYLSNENNADK 384

```

# RESULT 4 US-10-131-591A-23

```

; Sequence 23, Application US/10131591A
; Publication No. US20030059799A1
; GENERAL INFORMATION:
; APPLICANT: Nippon Zeon Co., Ltd.
; TITLE OF INVENTION: Modified DNA molecules, Recombinants and uses thereof
; FILE REFERENCE: J209
; CURRENT APPLICATION NUMBER: US/10/131, 591A
; PRIOR FILING DATE: 2002-08-15
; NUMBER OF SEQ ID NOS: 79

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 14, 2003, 10:21:13 ; Search time 28 Seconds  
(without alignments)  
2133.482 Million cell updates/sec

Title: US-09-901-572a-3  
Perfect score: 2324  
Sequence: 1 MHYFRNCIFFLIVLYGTN.....SSNNENADKIPGYRRPCTFL 456

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 492763 seqs, 131003257 residues

Total number of hits satisfying chosen parameters: 492763

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2317	99.7	456	US-09-147-052-2	Sequence 2, Appl1
2	2008	86.4	1086	US-09-147-052-4	Sequence 4, Appl1
3	1936.5	83.3	384	US-10-131-591A-12	Sequence 12, Appl1
4	1770	76.2	357	US-10-131-591A-23	Sequence 23, Appl1
5	320	13.8	62	US-10-131-591A-5	Sequence 5, Appl1
6	308	13.3	62	US-10-131-591A-6	Sequence 5249, Appl1
7	176.5	7.6	520	US-08-781-986A-5249	Sequence 2, Appl1
8	173	7.4	520	US-10-056-052-2	Sequence 5639, Appl1
9	170	7.3	2086	US-09-815-242-5639	Sequence 12610, A
10	169.5	7.3	5795	US-09-815-242-12610	Sequence 12996, A
11	159	6.8	6281	US-09-815-242-12996	Sequence 15885, Ap
12	159	6.8	1029	US-09-815-242-5885	Sequence 13083, A
13	159	6.8	1048	US-09-815-242-13083	Sequence 108, Ap
14	156	6.7	807	US-09-820-843A-108	Sequence 5835, Ap
15	156	6.7	2434	US-09-815-242-5835	Sequence 5835, Ap

16	154.5	6.6	2478	US-09-815-242-5816	Sequence 5816, Ap
17	154.5	6.6	2478	US-09-815-242-12967	Sequence 12967, A
18	153.5	6.6	3158	US-09-815-242-12611	Sequence 12611, A
19	147	6.3	1021	US-09-815-242-5471	Sequence 5471, Ap
20	147	6.3	1021	US-09-815-242-12544	Sequence 12544, A
21	146.5	6.3	837	US-09-815-242-5883	Sequence 5883, Ap
22	146.5	6.3	875	US-09-815-242-13080	Sequence 13080, A
23	143	6.2	1215	US-09-815-242-5908	Sequence 5908, A
24	143	6.2	1269	US-09-815-242-13113	Sequence 13113, A
25	143	6.2	2437	US-09-815-242-5834	Sequence 5834, A
26	140.5	6.0	596	US-09-946-374-243	Sequence 243, Ap
27	140.5	6.0	596	US-10-015-387A-243	Sequence 243, Ap
28	140.5	6.0	596	US-10-063-735-100	Sequence 243, Ap
29	140.5	6.0	596	US-10-006-130A-243	Sequence 243, Ap
30	140.5	6.0	596	US-10-199-672-310	Sequence 310, Ap
31	140.5	6.0	596	US-10-006-867-100	Sequence 310, Ap
32	140.5	6.0	596	US-10-052-586-310	Sequence 310, Ap
33	140.5	6.0	596	US-10-063-5847-100	Sequence 310, Ap
34	140.5	6.0	596	US-10-174-590-310	Sequence 310, Ap
35	140.5	6.0	596	US-10-176-758-310	Sequence 310, Ap
36	140.5	6.0	596	US-10-175-737-310	Sequence 310, Ap
37	140.5	6.0	596	US-10-063-616-100	Sequence 310, Ap
38	140.5	6.0	596	US-10-173-706-310	Sequence 310, Ap
39	140.5	6.0	596	US-10-175-728-310	Sequence 310, Ap
40	140.5	6.0	596	US-10-175-752-310	Sequence 310, Ap
41	140.5	6.0	596	US-10-176-482-310	Sequence 310, Ap
42	140.5	6.0	596	US-10-176-757-310	Sequence 310, Ap
43	140.5	6.0	596	US-10-176-913-310	Sequence 310, Ap
44	140.5	6.0	596	US-10-180-552-310	Sequence 310, Ap
45	140.5	6.0	596	US-10-180-557-310	Sequence 310, Ap

## ALIGNMENTS

RESULT 1  
US-09-147-052-2  
Sequence 2, Application US/09147052  
Patent No. US20010014335A1  
GENERAL INFORMATION:  
APPLICANT: SAITOH, Shuji  
APPLICANT: TSUZAKI, Yoshihara  
APPLICANT: YANAGIDA, No. US20010014335A1  
TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR,  
TITLE OF INVENTION: RECOMBINANT VIRUS, AND ITS USE  
FILE REFERENCE: 981167  
CURRENT FILING DATE: 1999-04-05  
CURRENT FILING DATE: 1999-04-05  
PRIOR APPLICATION NUMBER: JP 08-103548  
PRIOR FILING DATE: 1996-03-29  
PRIOR APPLICATION NUMBER: PCT/JP97/01084  
PRIOR FILING DATE: 1997-03-28  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 456  
TYPE: PRT  
ORGANISM: hybrid  
US-09-147-052-2  
Query Match  
Best Local Similarity 99.7%; Score 2317, DB 9, Length 456;  
Matches 454; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MHYFRNCIFFLIVLYGTN...SSNNENADKIPGYRRPCTFL 60  
1 MHYFRNCIFFLIVLYGTN...SSNNENADKIPGYRRPCTFL 60  
DB 61 EGGCSITKKANPNNGOTLEAARMELTDLANKAMTLASLOVAKIESLSAYSAE 120  
61 EGGCSITKKANPNNGOTLEAARMELTDLANKAMTLASLOVAKIESLSAYSAE 120  
QY 121 TVNNINLMTLEOLKRAKTNLESAINQANTDKTPNHNPNVENVKLLKTLLEGRATYLE 180